Migration Costs and Higher Education: Experimental Evidence from India

Merve Demirel*

Job Market Paper

November 5, 2026

Click here for the latest version.

Abstract

Opportunities are unevenly distributed across space. For people who live far from them, the non-pecuniary costs of migration, such as worries about safety, fear of loneliness, or anxiety about navigating an unknown city, determine whether they can access potentially transformative opportunities. Such costs are difficult to observe and quantify, limiting our understanding of their nature and of how to reduce them. This paper studies the role of migration costs in one of the most consequential decisions young people make: whether and where to pursue higher education. I ask whether migration costs hold prospective students back from migrating for college, and whether learning about student life in other cities can reduce perceived costs and facilitate migration. I follow 1,682 final-year secondary students across 52 schools in Maharashtra, India, over a full year as they finish school, apply, and enroll in higher education. First, I measure their perceived migration costs in an incentivized choice experiment and decompose them into beliefs about specific non-pecuniary outcomes at each destination. Next, I run a randomized controlled trial offering half of them a mentorship program that pairs each student with a university student who migrated for college. At baseline, the average student would forgo a 129% increase in expected post-graduation earnings rather than study away from home, an amount that far exceeds the returns students expect and rises steeply with distance. Their beliefs about the difficulty of moving, safety, parental approval, and happiness away from home account for a substantial portion of perceived migration costs. Mentorship increases a pre-specified educational migration index by 0.16 standard deviations, driven by short-distance moves made by male students. Consistent with the stated-preference evidence, these effects are concentrated on the margin where migration costs are smallest, suggesting these costs pose a major obstacle to educational mobility.

*PhD student, Institute for International Economic Studies (IIES), Stockholm University. Email: merve.demirel@iies.su.se. I am grateful to Johannes Haushofer, Anna Tompsett, Tessa Bold, and Laia Navarro-Sola for their guidance and support, and to Livia Alfonsi, Ingvild Almås, Toman Barsbai, Jakob Beuschlein, Palaash Bhargava, Dominik Biesalski, Monir Bounadi, Nina Buchmann, Cristina Clerici, Mitch Downey, Alex Eble, Prashikdivya Gajbhiye, Robert Garlick, Anahita Karandikar, Vera Kågström, Martin Koenen, Michael Kremer, Louis-Pierre Lepage, Francesco Loiacono, David McKenzie, Arash Nekoei, Nicklas Nordfors, Thore Petersen, Amalia Repele, Sandra Sequeira, Caterina Soto-Vieira, Sneha Subramanian, Bertil Tungodden, and audiences at the ASWEDE conference, the Oxford Development Economics Workshop, Columbia University, Stockholm University, the University of Groningen, and Tilburg University for valuable feedback and advice. Tejas Tule and Sahil Pawar provided excellent research assistance. The trial is registered in the AEA RCT Registry (AEARCTR-0016735), and pre-analysis plans for the midline and endline surveys were filed before the respective data were collected. The study received ethical approval from the DAI Research & Advisory Services Institutional Review Board (approval no. PN-MH-01/2025-2026) and the Swedish Ethical Review Authority (Etikprövningsmyndigheten, approval no. 2025-04262-01). I gratefully acknowledge funding from the Royal Swedish Academy of Sciences, the Weiss Fund for Research in Development Economics, the Mannerfelt Foundation, and SurveyCTO.

1 Introduction

Large gaps in economic opportunity across places in developing countries persist alongside low rates of migration (Gollin et al., 2014; Lagakos, 2020). A large literature attributes this to a disutility of migration itself: it is costly to leave home (Sjaastad, 1962; Kennan and Walker, 2011; Bryan and Morten, 2019; Lagakos et al., 2023). Less is known about what these costs consist of, or whether they can be changed.

For many people, the first migration decision comes with the choice of where to study for higher education, which largely determines both the quality of education they receive and the labor market they enter, with potentially large economic gains that persist over one's lifetime (Munshi and Rosenzweig, 2016; Sekhri, 2020). Yet higher-education opportunities are unevenly distributed across space, accessible only to those who are able to migrate. Costs of migration are therefore barriers to higher education, and the resulting losses accrue throughout a working life. How large are these costs, what do they consist of, and can they be reduced enough to change whether and where students choose to study?

I study these questions in the context of college choice in Maharashtra, India, where students decide whether and where to continue their education. I recruit 1,682 high-school students in two districts of Maharashtra to a series of surveys and experiments as they complete their last year of schooling, apply to, and enroll in higher education. The median student in my sample expects that studying at a college away from home to lead to higher earnings net of living costs and could enter and pay for one, yet 56% name their home as their most preferred study destination. Whether this represents students' preferences over locations, or high costs of migration determines how we should understand spatial gaps: if the unwillingness to move reflects informed preferences, the gap consists of compensating differentials. Otherwise, the same gap can be characterized as a barrier to potentially transformative educational opportunities. The question then becomes whether there is room for policy interventions to lower these migration costs and help students access opportunity.

These explanations cannot be separated based on choice data alone: the same decision to stay home is consistent with a student who knows what life away from home would be like and dislikes it, and with one who has fears and uncertainty about migration, but would actually thrive if they took the leap (Manski, 2004). To separate the two, I proceed in two steps. First, in an incentivized choice experiment, I measure perceived migration costs directly and ask how much of these costs can be accounted for by beliefs about life at each destination. Second, in a randomized controlled trial, I evalu-

ate a mentorship program that connects these high-school students with college students who have left home to study, to test whether access to the information such a contact can provide affects perceived migration costs and students' realized application, enrollment, and migration decisions.

To measure how large perceived migration costs are and what they consist of, I implement a discrete choice experiment that isolates the disutility associated with moving from home to other destinations in the region. I present students with pairs of hypothetical colleges located at home or in real destinations and ask them to choose between the two, assuming admission to both. Hypothetical colleges allow me to vary a college's location independently of its other attributes, which are correlated with location among the colleges students can actually apply to, and estimate the causal effect of moving away from home to study on stated choices. Every college at a destination may not be feasible for every student, so the design holds feasibility and pecuniary aspects of moving fixed, leaving the two colleges to differ only in location and salary. The magnitude of salary gains a student forgoes to study at home therefore reveal the non-pecuniary costs of migration. Importantly, the survey doubles as the application to the mentorship program. Students are told that a lottery decides who receives a mentor and that their answers are used to choose which mentor they receive. The more truthful they answer, the better of a mentor match they can receive, so they have an incentive to answer truthfully.

I find that the average student would forgo an increase that more than doubles their expected post-graduation earnings rather than study away from home. Put differently, even when students are choosing between two equally attainable and affordable colleges, the non-pecuniary disutility of migration is large. The disutility increases with distance, and the earnings increase students would require is several times larger for the farthest destination than for the nearest city. The required increase far exceeds the pecuniary returns students expect from studying away. These high implied costs of migration make it unsurprising that so many prefer to and plan to study at home.

The migration costs estimated by the discrete choice experiment are subjective and multidimensional. To unpack what the estimates represent, I also elicit each student's beliefs about life in the same destinations, following the rich literature on elicitation of subjective expectations (Jensen, 2010; Wiswall and Zafar, 2015; Giustinelli, 2023), and estimate how much of the migration costs these beliefs account for. I focus on dimensions related to emotional wellbeing, social ties, and the practical experience of living and studying at each destination. Students expect to be worse off away than at home on every non-pecuniary outcome I elicit: they expect to be less happy and less safe, and they rate their chances of experiencing mental difficulties higher. They also anticipate

practical difficulties when moving far, higher parental and societal disapproval, fewer close friends, and less time with family. Expected declines grow with distance, as do the migration costs I infer from the choice experiments. Conditioning on each student's own beliefs about a destination attenuates the estimated migration costs by more than half, with beliefs about safety and the difficulty of moving accounting for the largest share of this reduction. Much of what looks like attachment to home therefore appears to reflect these beliefs, and how we should interpret migration costs depends on whether the beliefs are accurate.

Throughout, I find that men and women face very different trade-offs when making migration decisions. Migration costs are larger for women, who would require about three times the earnings increase that men would (227% for women versus 72% for men), albeit from a lower baseline of expected salary. Women also expect larger declines than men in safety, parental approval, and reputation back home if they were to move. Consistent with this, these beliefs account for a larger share of women's perceived migration costs than of men's.

Are these beliefs accurate? They may not be, since students have few sources of information to verify their beliefs against. At baseline, many report no source of college advice, and knowledge of colleges and destinations is limited, most of all for the major education hubs. Existing social networks appear to matter for both knowledge and beliefs: students who have a contact in a destination know more about colleges and student life there. They also hold more favorable beliefs about studying there, so much so that knowing someone at a destination seems to make a distant destination look similar to a close one: connected students' beliefs about a destination eleven hours' drive away resemble those of students without a contact about a destination two hours away. This pattern may arise because a contact in a distant destination lowers the costs of moving, as both theory and empirical evidence on pre-existing social networks (Carrington et al., 1996; Munshi, 2020; Sahai and Bailey, 2022; Koenen and Johnston, 2024; Blumenstock et al., 2025), as well as two studies that exploit chance encounters suggest (Battiston, 2018; Green, 2024), highlighting that even brief interactions can have large effects when they occur in high-stakes decision settings. Students who have a contact, however, may differ in other ways from those who do not, so this pattern alone cannot tell us whether an additional connection can reduce perceived costs and facilitate migration.

Against this backdrop, I design and implement an intervention that gives students access to the people best placed to inform beliefs about student life away from home: college students who themselves left home to study in one of the major education hubs, Pune and Mumbai, which are also common destinations for students in my sample who

leave their district. I randomly offer half of the students in my sample a place in *"Campus Trail"*, a phone-based mentorship program that matches each student with a mentor of the same gender and field of study. In the months before college application deadlines, the mentors and mentees speak by phone about college and major choice, the practicalities of moving, and the mentor's own experience of studying away from home, and mentors offer students help researching their options and deciding among them. The intervention builds on evidence that mentors and role models influence schooling and early career choices (Beaman et al., 2012; Bettinger et al., 2012; Hoxby and Turner, 2013; Castleman and Page, 2015; Carrell and Sacerdote, 2017; Breda et al., 2023; Alfonsi et al., 2024; Riley, 2024), and a consistent lesson from this work is that students facing complex decisions need personalized support, including information, encouragement, and practical help with various application procedures. The program extends that lesson to a distinct margin of college access: barriers generated and amplified by spatial distance. I estimate the intent-to-treat effects of the offer on realized application, enrollment, and migration decisions at endline, and on destination-specific beliefs.

Conversations with a migrant peer shift students' applications and realized moves. At endline, the offer raises a pre-specified migration index that combines indicators of continuing in higher education with indicators of migration by 0.16 standard deviations. Each of the index's six components rises, and two do so significantly: treated students are three percentage points more likely to apply for further study and five points more likely to be enrolled. Given high base rates, these are sizable effects, reducing the share of students who do not apply by about a third and the share who do not enroll by about a fifth. Applications to colleges away from home rise by four percentage points and moves away from home by three, both increases of about 12 percent, but neither is statistically significant. Consistent with the stated-preference evidence, the effect is driven by moves over shorter distances, where perceived migration costs are smallest. Treated students are five percentage points more likely to apply to study elsewhere in their home district and three points more likely to move there, from bases of 18 and 10 percent. Applications and moves to the major education hubs, where the mentors study, do not increase.

Men and women respond differently to the mentorship program. Men's treatment effects are concentrated on the intensive margin. Mentorship does not change whether they continue to higher education, but it does change whether they migrate to do so. Women's treatment effects are concentrated on the extensive margin: treated women are more likely to apply for and enroll in higher education, and the patterns are suggestive of a displacement effect of their applications from farther destinations toward nearer

ones.

Mentorship could have changed behavior through students' expectations about pecuniary and non-pecuniary outcomes at destinations, or through encouragement, support, and practical help. I find most supporting evidence for the mentorship operating through expectations, which I elicit at endline for home and for the closest city within or outside of students' district. For men, mentorship mostly changes how they perceive the pecuniary outcomes at different destinations: treated men expect lower living costs and higher earnings after graduation, relative to home. Their expectations about non-pecuniary outcomes do not change.

For women, the estimates suggest that talking to a mentor makes studying away look worse: expected returns to studying away decrease and, if anything, beliefs about life in the closest major education hub worsen relative to home. I read this as consistent with the large migration costs students perceive at baseline being broadly accurate rather than misperceived. If these beliefs were largely driven by misperceptions about life away from home, contact with someone who has moved away to study could correct them. I find evidence to the contrary, which points to real migration costs, large enough to not be affected by a light-touch peer mentorship intervention, as a key barrier to educational mobility.

My findings have important implications for policy. A natural supply-side response to spatial inequalities in educational opportunity is to bring education closer to students by opening new colleges. A college within commuting distance may raise local enrollment[1], but if migration costs still bind, its graduates will largely enter the labor market where they studied (Conzelmann et al., 2025). Unless new colleges match the educational and labor market opportunities of existing ones further away, such policies may fall short of delivering the full transformative effects of education. If so, there is a case for moving students to education rather than education to students. Consistent with this, Ghose (2026) estimates that in India, lowering the cost of moving for education would raise welfare by more than lowering the cost of moving for work, because the student who moves for college need not move again for a job. My study evaluates a demand-side intervention that aims to move students to educational opportunities, and shows that migration costs are a major obstacle to these moves. The results of the two experiments in my study suggest that encouraging educational migration at scale likely means lowering migration costs at the destination, for example through safe and affordable student housing, practical help with moving, or support that continues after

[1] A large literature evaluating these policies finds that opening colleges raises enrollment (Katzkowicz et al., 2026), especially among women (Elsayed and Shirshikova, 2023; Caner et al., 2024; Sekhri et al., 2024).

students arrive.

My first contribution is to the literature examining frictions to internal migration. This literature consistently finds that the migration costs needed to rationalize observed migration are far larger than the observable monetary costs of a move, suggesting that non-pecuniary costs play a large role, with consequences for productivity, human capital accumulation and inequality (Sjaastad, 1962; Kennan and Walker, 2011; Bryan and Morten, 2019; Imbert and Papp, 2020; Fu et al., 2022; Lagakos et al., 2023; Fabre, 2024; Ishimaru, 2025; Ghose, 2026). Because non-pecuniary costs are hard to observe, most of these estimates are catch-all terms, inferred from migration choices assumed to be optimal or elicited as a single preference for home (Batista and McKenzie, 2023; Batista et al., 2025).[2] Other studies estimate the value particular attributes of a destination, such as housing quality, crime, or proximity to family (Baláž et al., 2016; Koşar et al., 2022; Lagakos et al., 2023). I bring the two approaches together for a large cohort facing an important migration decision: I measure the full non-pecuniary cost of leaving home, show that it far exceeds the pecuniary returns students expect, and decompose it into each student's expectations about specific outcomes at the destinations they face. Much of what looks like attachment to home reflects these expectations: students prefer home not only because it is home, but because they expect to be safer and happier there, and expect moving away to be hard.

There are reasons to doubt that these perceived costs are accurate or fixed (McKenzie, 2024). Whether they are accurate matters for policy: if students overestimate how hard life away from home would be, better information could be a cheap way to help students access educational opportunity. Potential migrants can be misinformed about other aspects of a potential move, as is often the case with what they would earn at the destination (McKenzie et al., 2013; Baseler, 2023). Attachment to home may also reflect habit (Sarvimäki et al., 2022) or unfamiliarity that fades with experience (Lagakos et al., 2023; Dean et al., 2026). I test whether contact with a peer who has moved away to study changes these perceived costs. Such contact does not improve students' expectations about life away from home, suggesting that high migration costs are unlikely to be driven by misperceptions.

My study relates to a literature that experimentally varies the information available to potential migrants (Bryan et al., 2014; Shrestha, 2020; Baseler, 2023; Frohnweiler et al., 2024; Bhatiya et al., 2026). These studies primarily provide information about earnings and employment opportunities at destination, often shifting beliefs but producing

[2] An exception is Bhatiya et al. (2026), who estimate migration costs from real choices with measured and experimentally shifted beliefs: correcting rural job seekers' beliefs about where the jobs offered by an Indian training and placement program are located changes whether they enroll, implying that they require 50 percent higher pay to work outside their home state than within it.

mixed effects on migration behavior. A smaller strand of this literature uses exposure to lived migration experience to inform potential migrants through firsthand accounts and advice from migrants. Closest to my study, Barnett-Howell et al. (2026) study rural households in Kenya who, as in my setting, migrate little despite large returns to moving. Information about urban earnings, which these households underestimate, raises migration and household income, both on its own and when paired with access to a mentor at the destination. By contrast, Mesplé-Somps and Nilsson (2023) expose potential migrants to documentaries portraying migrant role models, raising aspirations but not the desire to migrate. Unlike mentorship, these documentaries provide exposure to lived migration experience without opportunities for personalized exchange.[3] My study differs from this work in the decision it studies and in the timing and breadth of the contact it offers. First, it examines the choice of where to study: a one-time, high-stakes decision made at the start of adult life that shapes where students go on to live and work. Second, mentors talk with students while they are making this decision, and their conversations cover both the academic aspects of the decision, such as courses, colleges, and the careers they lead to, and the everyday considerations of living away from home: finding a hostel, managing money, getting around an unfamiliar city, making friends, and handling parents' worries about a child living far away. Together, these features let me ask how contact with someone who has already moved away to study influences a consequential decision.

I also contribute to work on gendered mobility. A consistent finding in this literature is that women give up more educational and economic opportunities than men to stay close to home or to travel safely. Faced with similar options, women are more willing than men to accept a lower-quality college or a lower-paying job in exchange for a shorter or safer journey (Borker, 2021; Le Barbanchon et al., 2021; Alba-Vivar, 2025), and many forgo valuable training altogether when it means leaving their village (Cheema et al., 2026). Experiments show that women's choices respond strongly when their daily commute becomes faster or safer: providing bicycles to schoolgirls, women-only transport, or other women to travel with raises girls' secondary enrollment and women's job search (Muralidharan and Prakash, 2017; Garlick et al., 2025; Kapoor and Gade, 2025). Much less is known about the gendered patterns of migration, a longer-term and higher-stakes mobility decision, and less still about how to ease the frictions that shape these patterns. The few studies of the decision to migrate examines higher education choices, and shows that women apply less often than comparable men to institutions far from home (Farré

[3] A closely related but distinct literature studies information interventions concerning irregular migration from West Africa to Europe, often combining migrant testimonials with factual information. These interventions can alter perceptions and stated intentions (Tjaden and Dunsch, 2021; Bah et al., 2023), with some evidence that they also reduce actual migration (Battiston et al., 2024).

and Ortega, 2024; Han et al., 2025). Demirel et al. (2026) also study students in India as they enter higher education. Estimating a choice model among applicants to elite engineering programs, they find that women weigh distance from home substantially more heavily than men, enough to account for about a quarter of the gender gap in the value of the programs they can attain. My findings suggest that these gaps partially reflect higher perceived non-pecuniary costs of moving away from home.

Finally, I contribute to the literature on non-pecuniary returns to educational investments, which studies how young people's expectations about outcomes beyond earnings shape whether, what, and where they study (Boneva and Rauh, 2017; Delavande and Zafar, 2019; Clerici et al., 2024). My findings echo this literature's conclusion that non-pecuniary aspects, such as enjoying one's studies, a school's ideology, or family approval, can be far more important than labor market returns. I show that this holds for deciding where to study as well: many students expect to earn more if they study away, yet their beliefs about what daily life in a different city would look like account for much of their preference to stay.

The remainder of the paper proceeds as follows. Section 2 describes the setting, sample, and data. Section 3 outlines the conceptual framework. Section 4 documents the size of migration costs, the beliefs behind them, and the information environment in which those beliefs form. Section 5 describes the mentorship experiment, and Section 6 describes the empirical strategy. Section 7 presents the treatment effects of the intervention. Section 8 concludes.

2 Setting and sample

2.1 Higher education and migration in Maharashtra

The study follows 12th-grade students sampled across 52 schools in two districts in Maharashtra (Ahilyanagar and Sangli), as they decide where to continue their education. Students can choose to continue their education in the sub-district in which they are observed at baseline,[4] which I refer to as *home* throughout the paper, or *away*, meaning any destination outside it. Within the students' districts, district capitals (approximately two hours' drive away on average for the sample) offer the largest number of institutions and opportunities; outside the districts, there are many options students can consider throughout the state. The study focuses on understanding how students view a total of five options when it comes to their study choices: their home, their district capital, and

[4] A sub-district (*taluka*) comprises multiple villages and towns. I define a student's home as the sub-district in which their school is located and treat it as the student's local geographic, labor-market, educational, and social environment.

the three largest higher-education hubs outside the study districts (Pune, Mumbai, and Nagpur). Figure 1 shows the geography of the study sites and the destinations, as well as the distances between them.

Of the three hubs, Pune and Mumbai[5] represent the state's two largest higher-education and labor markets, and the mentors are selected from among college students in these regions. According to the All India Survey on Higher Education (AISHE) data from 2022, these two regions together host a quarter of Maharashtra's universities and affiliated colleges and a third of the students enrolled in higher education. They also receive the largest share of educational migrants within the state (28 % in the 2011 Census).[6] They are four to eight hours' drive away, and Pune is the most common out-of-district study destination in my sample (6 percent of control group students study there). Nagpur presents a third and farthest option, which offers a considerable number of colleges but not the same access to a large labor market, and I consider it only in the baseline survey. All destinations in the study share a common language (Marathi), a common state administration, and the same eligibility rules for state reservations.

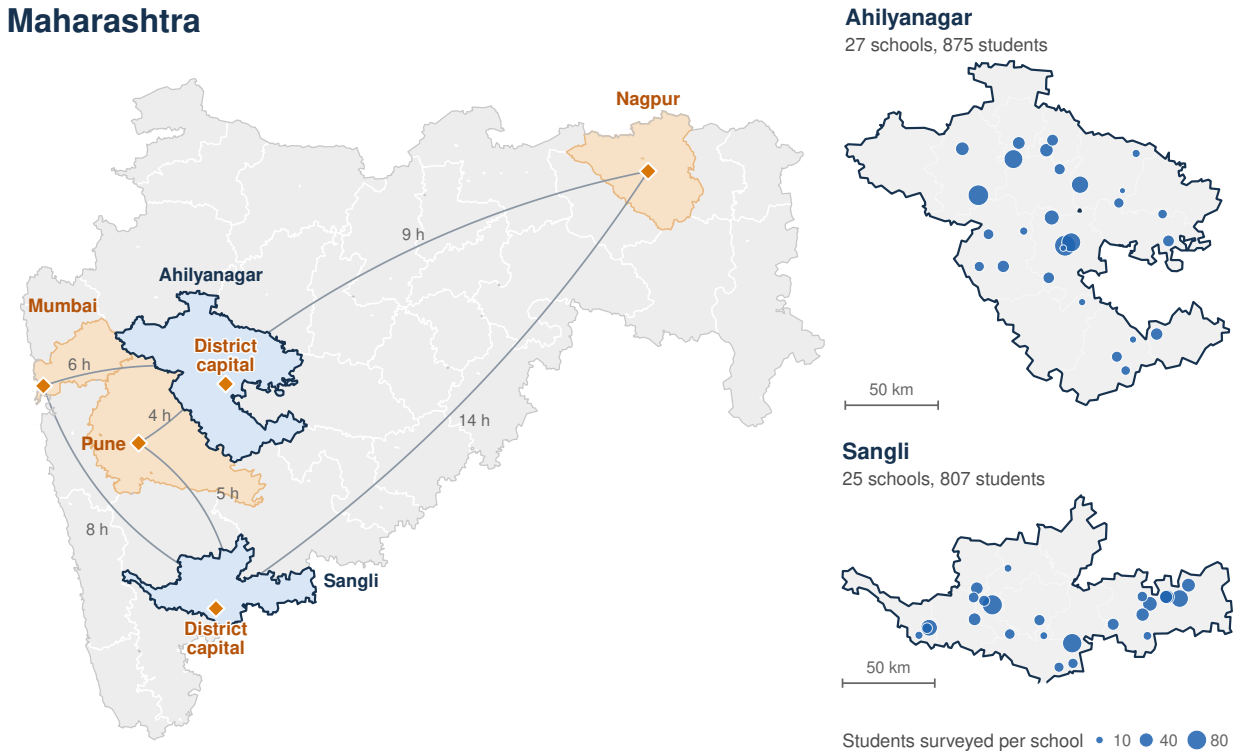
For a prospective college student in the two study districts, cities a few hours of travel away therefore offer drastically wider access to higher education. The fact that the number of options increases as the search area expands is unsurprising; Appendix Figures A2 and A3 show that this is also true conditional on academic achievement, throughout the distribution. Among the institutions that admit through centralized exams, even the highest-achieving student can reach on average only about 90 engineering-program seats within their *home*, around 1,000 elsewhere in the district, and roughly 21,500 outside the district. When adjusted for the number of other qualified candidates who can compete for the same seats, these numbers are smaller, but the advantage remains for most of the score distribution. On top of expanding the choice set of the students, colleges in the two regions also represent more sought-after alternatives, as evidenced by their admissions cutoffs. In the centralized admissions systems, seats are allocated by entrance-exam score: a program's cutoff is high when many high-scoring applicants rank it above other programs, therefore cutoffs provide a revealed-preference measure of college quality. The median engineering program in Pune and the Mumbai region has a cutoff at the 82nd percentile of the entrance exam, compared with the 46th percentile in the study districts (Appendix Figure A5), and admission to such colleges carries substantial labor-market returns in India (Sekhri, 2020).

In addition to more educational opportunities, Pune and the Mumbai region offer

[5]Mumbai refers to the Mumbai region: Mumbai, Mumbai Suburban, and Thane districts.

[6]Census of India 2011, Table D-3 (migrants by place of last residence and reason for migration): education migrants whose last residence was elsewhere in Maharashtra; Pune district and the Mumbai region together account for 182,968 of 648,596.

Figure 1: Study sites and destinations



Notes: The left panel shows Maharashtra, with the two study districts (Ahilyanagar and Sangli) in blue and the destinations the survey asks about in orange, including the capital of each study district. The average driving time to the district capital, among students who do not live there, is 2 hours in both districts.

access to larger and higher-paying labor markets. In India, urban wages are far above rural wages (Chauvin et al., 2017), so, even holding college quality fixed, studying in these destinations may help students access better jobs through internships, job-search networks, or more information about local employers. Returns to migration are hard to estimate for my study sample because most household surveys are not representative at the district level. Several sources nonetheless point in the same direction. Per-capita incomes in Pune and the Mumbai region are 60–90% higher than in the study districts,[7] and in the India Human Development Survey, per-capita incomes in these two regions are 106% higher and consumption 56% higher than in the rest of the state after adjusting for spatial price differences (Appendix Table A1). Satellite data on nighttime lights, a proxy for economic activity, show a similar gap.[8] These observational differences reflect

[7]Source: Directorate of Economics and Statistics, Planning Department, Government of Maharashtra, *District Domestic Product of Maharashtra 2011–12 to 2023–24* (`https://mahades.maharashtra.gov.in`).

[8]Light per resident in the two regions is 1.8 to 2.2 times that in the study districts in every year from 2014 to 2025. Radiance is from the VIIRS satellite sensor, cleaned and aggregated to 2011 district boundaries by Patnaik and Mahalingam (2026), summed over the year and divided by 2011 Census population. On 2011 boundaries, Thane includes Palghar.

both selection and place effects, so they should not be interpreted as the causal return a marginal student would obtain by migrating; they do, however, indicate the magnitude of the gap a prospective educational migrant is weighing.

Despite these differences, migration for education in Maharashtra is uncommon and falls steeply with distance: according to the 2011 Census, only 1.35 % of within-state migrants moved for education, and 57.5 % of those educational migrants stayed within their own district.[9] Low mobility despite large spatial wage gaps is a general feature of India (Munshi and Rosenzweig, 2016), and India is the least urbanized of the world's ten largest economies (Chandrasekhar and Sharma, 2015). A fragmented admissions system, with multiple entrance exams that vary by program and region and with state- and education-region-level quotas that reserve seats for local applicants, reinforces the incentive to stay close to home. In this study's sample, 9.0% of graduates in the control group migrate within their district for education, 11.8% to another district, and virtually none across state borders.

2.2 Sample and data

The sample consists of 1,682 final-year secondary-school students across 52 schools who signed up for a mentorship program with college students in Pune and Mumbai when it was presented in their school (Section 5). Sixty percent of students are female, 68% study the science stream, 36% belong to the general (non-reserved) caste category, and the median student is seventeen. Seventy percent of fathers and about 50% of mothers have completed class 10, higher shares than in other Maharashtra households with a teenager who has completed class 10, although fewer parents hold a degree.[10] Study-sample schools are more likely to be rural than other schools in the study districts, and they have fewer teachers and classrooms per student (Appendix Table B1).

The data come from three student survey waves, a parent survey, and administrative school records (Appendix C shows the study timeline). The baseline survey, administered on tablets in school in August and September 2025, precedes the intervention and covers household and demographic characteristics, academic background, measures of information and search, and trait batteries covering self-efficacy, gender attitudes, aspirations, and social desirability. Importantly, it also administers the discrete choice

[9]The patterns are similar at the national level. While 38 % of the Indian population moved within the country, only 1.1 % of all migrants moved for education, compared to 50 % for marriage and 10.4 % for employment or business. Although 20.8 % of young Indians are in higher education (Ministry of Human Resource Development, 2013), only 3.3 % migrate for education. Among those who do move, 53.5 % stay within their own district and only 13.9 % cross state lines. Migration is also strongly gendered; Section 4.2 returns to gender heterogeneity in the context of migration costs.

[10]In households in Maharashtra with a member aged 16–18 who has completed class 10, 60 percent of fathers and 46 percent of mothers have completed class 10, and 16 and 11 percent hold a degree (Household Consumption Expenditure Survey 2023–24, which does not record school enrollment). In the sample, 8 percent of fathers and 4 percent of mothers hold a degree.

experiment and college rating exercise as well as the elicitation of destination-specific beliefs, described in detail in Section 4.1, for three randomly selected destinations out of the five described earlier. Because the baseline survey doubles as the application to the program and students are told that their answers are used to match them with a mentor, students have an incentive to answer truthfully (Section 4.1). The midline survey, administered in person in schools between November 2025 and January 2026, before entrance exams and admissions, re-administers the discrete choice experiment, the college rating exercise, and the belief, information, and trait questions. The endline survey, conducted by phone from July 2026 as admissions conclude, measures realized applications, enrollment, and migration. Alongside the student baseline, I also survey one parent or guardian per student by phone, reaching a parent for 792 of the 1,682 students (47%). School characteristics come from the Unified District Information System for Education (UDISE+), and school-level results come from the 2026 twelfth-grade board exam.

Students in the sample are preparing for the Higher Secondary Certificate (HSC) board exam at the end of twelfth grade, which is required for graduation and serves as one admission route: many institutions admit on the basis of HSC marks alone. Admission to many programs also requires one or more entrance exams (Appendix A). At endline, 75% of science students and less than 10% of commerce or arts students report having sat at least one, almost always the state entrance exam. Vocational and diploma programs have their own admission systems. Coaching, either as after-school classes during higher secondary school or as full-time classes afterward, is a standard part of preparation: 60% of students report already attending or planning to attend coaching classes at baseline.[11]

Three features of the sample matter for interpreting the study. First, the students in my sample are aiming to study further and are interested in speaking with a peer mentor who is studying in a college in Pune or Mumbai. They signed up to speak with such a mentor, and 95 percent say at baseline that they would like to continue their education after twelfth grade. These plans carry through to endline, when 91 percent of control students have applied or plan to apply for further study (Table 3). The sampling strategy leaves out students who do not plan to continue their education, but it does not condition on migration plans, so the sample is not limited to students who have already decided to study away. Second, educational migration is academically feasible for the average student in the sample. The typical study-sample student lies around the 25th percentile of students placing into colleges in Pune and the Mumbai region through the

[11] Twenty-four percent already attend coaching, the same share as class 11–12 students in rural Maharashtra (32 percent statewide; Comprehensive Modular Survey: Education 2025).

state entrance exam, and Appendix Figure B1 shows that HSC results in study-sample schools lag only slightly behind district and state averages. Third, studying away is financially feasible for most students in the sample: about two-thirds of students (81 percent of parents) report that their family could afford the median fees of a government or aided college in urban Maharashtra ($153 a year). Their stated monthly living budgets are comparable to what a representative student in Pune spends (Mangal, 2024), although at the lower end of those in the mentor applicant pool (Appendix D).[12] Together, these facts suggest that the average student could gain admission to and pay for some colleges away from home. Therefore, in this paper, feasibility refers to destinations as a whole, not to individual colleges. Students in the sample are therefore selected to be on the relevant margin for the study because they intend to continue their education and studying away from home is broadly an attainable alternative.

When asked about their expectations about returns to studying away, most students expect substantial pecuniary gains from studying away from home. At midline, control students expect salaries 32% higher away than at home on average, and 65 % expect a positive return net of living costs from at least one of the elicited destinations.[13] This is a lower bound on the share who expect a net positive return to studying somewhere away from home. Among them, the median stated differential is roughly $15,100 in present-value terms, with a long right tail (mean $33,000); over all control students, the median differential is about $6,300.[14] Yet when asked directly for their most preferred option, 44 % of these same students name home, and only 35 % name a destination for which they expect a positive return.[15] The rest of the paper asks whether the gap between students' stated choices and their expected pecuniary returns reflects large non-pecuniary migration costs, and whether talking to a peer who has moved away to study

[12]Median annual course fees of undergraduates in government or aided colleges in urban Maharashtra in the 75th round of the National Sample Survey (Household Social Consumption: Education, 2017–18), converted to 2024–25 prices with the combined consumer price index. Students report the most their family could afford for yearly college fees in brackets, and the threshold falls inside the lowest bracket (below $267), so the brackets alone bound the student share between 41 and 100 percent; the two-thirds figure uses the distribution of parents' exact answers within that bracket. Parents report exact amounts. USD at 93.75 INR.

[13]The baseline city module elicited salaries in brackets of about $1,070 and measured home at the district level, which compresses the differentials. I therefore report midline answers, which record exact amounts. At baseline, students expect salaries 16% higher away than at home on average, 56 % of all students expect a positive return from at least one of the two elicited cities (median differential among them about $15,000), and 54 % of these students name home as their most preferred option. With bracket midpoints in place of exact amounts, the midline share falls from 65 to 61 %.

[14]Midline control students (N = 614); treated students' midline answers may reflect the intervention. The midline module elicits the district capital, Pune, and Mumbai, and measures home at the sub-district level. For each elicited destination, the pecuniary differential combines the stated monthly salary and monthly living-cost gaps relative to home. The present value assumes the living-cost differential is borne over four years of study and the salary differential earned over the following forty years, discounted at 5 % annually; each student's best destination is used. The positive-return share is 64 % when the salary differential is weighted by the student's stated probability of working full-time at age 25. In a second conservative check, expected salary after studying away is $p_{ic}^{\text{stay}} w_{ic} + (1 - p_{ic}^{\text{stay}}) w_{ih}$, where p_{ic}^{stay} is the student's destination-specific probability of remaining there after graduation; the positive-return share is 59 %. Applying both probability adjustments yields 57 %. Restricted to Pune and Mumbai, the positive-return share is 60 %.

[15]Most preferred option from the midline main survey's direct ranking question. A ratings-based alternative (the strictly highest-rated of the four elicited options on the 0–10 scale) gives 30 %.

can lower these costs and change students' college choices.

3 Conceptual framework

I conceptualize students' college migration decisions as a destination-choice problem. Student i considers a set of destinations c and chooses the option that maximizes utility:

$$U_{ic} = f(V_{ic}, N_{ic}).$$

Here V_{ic} is the student's perceived net *pecuniary value* of attending college in destination c: the present value of expected post-graduation earnings net of expected pecuniary costs, including tuition fees net of scholarships, travel, housing, and living expenses. N_{ic} is the student's perceived net *non-pecuniary value* of attending college in c, bundling safety, parental approval, social ties, mental wellbeing, and other lived-experience factors. Utility is increasing in both arguments. The options are destinations, and each destination offers a range of colleges. Section 2 shows that most students could enter and pay for some college away from home, so the framework abstracts from admission and budget constraints.

The objects V_{ic} and N_{ic} are subjective. They summarize the values that enter student i's decision, combining beliefs about each destination with the weights students place on those expected conditions. Two students may therefore assign different V_{ic} or N_{ic} to the same destination: they may hold different beliefs about safety, costs, or returns, or they may place different weight on the same expected conditions.

Home is the default option and the natural reference point for the migration decision. A student migrates from home h to another destination c when

$$U_{ic} > U_{ih}.$$

The relevant comparison is therefore the destination-home utility gap:

$$\Delta U_{ic} \equiv U_{ic} - U_{ih},$$

with corresponding pecuniary and non-pecuniary gaps

$$\Delta V_{ic} \equiv V_{ic} - V_{ih}, \qquad \Delta N_{ic} \equiv N_{ic} - N_{ih}.$$

I call ΔN_{ic} the *non-pecuniary value differential*: the non-pecuniary value student i attaches to studying in c rather than at home. A positive differential means the student expects

to be better off away on non-pecuniary grounds; a negative differential is a disutility of migrating that the pecuniary gain ΔV_{ic} must overcome. Section 4 measures this differential and finds it negative for the average student at every destination; I refer to its magnitude, expressed as the earnings increase that compensates for it, as *migration costs*.

A student may expect higher net pecuniary value from studying away from home, $\Delta V_{ic} > 0$, but still prefer home if the non-pecuniary gap is sufficiently negative. Conversely, a student may migrate because the pecuniary gain is large, because the non-pecuniary cost is small, or because the destination dominates home on both margins. Migration selection may therefore reflect high expected pecuniary returns away from home, low perceived migration costs, low sensitivity to those costs, or a combination of these forces.

The framework implies two conditions for the intervention to change observed migration choices. First, the intervention must raise the utility of at least one non-home destination relative to home:

$$\Delta U_{ic}^{post} > \Delta U_{ic}^{pre}.$$

Because utility is increasing in V and N, this can occur if the intervention raises the perceived or effective pecuniary value of the destination relative to home, ΔV_{ic}, or the perceived or effective non-pecuniary value of the destination relative to home, ΔN_{ic}.

Second, the shift must be large enough to reverse the student's preferred option. For a student who prefers home at baseline, this requires

$$\Delta U_{ic}^{pre} < 0 < \Delta U_{ic}^{post}.$$

Equivalently, the treatment-induced shift must exceed the student's baseline distance to the migration margin:

$$\Delta U_{ic}^{post} - \Delta U_{ic}^{pre} > -\Delta U_{ic}^{pre}.$$

The intervention may therefore change beliefs or reduce costs for many students, but observed migration changes only for students whose destination ranking is reversed.

Mechanisms The intervention can satisfy the first condition through several channels. First, peer conversations may change students' beliefs about destination-specific returns or costs. For example, students may revise their expectations about safety, loneliness, commuting, parental acceptance, living conditions, or the practical difficulty of navigating a destination. They may also learn about scholarships and educational loans, which change the expected cost of studying away and how it can be financed.

Second, peer conversations may reduce effective migration costs directly. A peer may help identify safe housing, explain how to move around the city, provide contacts, or make the destination easier to navigate. In this case, the intervention changes not only beliefs about migration costs but the costs themselves.

Third, the intervention may change the salience or perceived importance of particular costs, making the same expected difficulty feel more manageable. A change in ΔN_{ic}, for example, may therefore reflect belief updating about a fixed destination environment, a real reduction in migration costs through peer support, or a change in how strongly the student weights a given expected condition. The three channels are observationally equivalent in my data, because the survey asks students how they expect to experience a destination, for example how safe they would feel there. Section 5 returns to how the intervention addresses each of these mechanisms.

Migration responses The ranking-reversal condition clarifies when changes in the migration calculus translate into observed migration. Students with ΔU_{ic}^{pre} close to zero require only a small shift in perceived or effective returns and costs for a non-home destination to become preferred. Students with large negative values of ΔU_{ic}^{pre} are unlikely to migrate unless the intervention produces a large shift. Students with large positive values of ΔU_{ic}^{pre} would migrate even without the intervention. Treatment effects on observed migration should therefore be concentrated among students close to the migration margin.

The framework motivates the empirical analysis that follows. Section 4 quantifies the non-pecuniary gap ΔN_{ic} and the beliefs behind it. Section 7 then estimates whether the intervention changes realized migration at endline, where the induced moves go, and which beliefs underlying ΔN_{ic} move, and for whom.

4 The non-pecuniary cost of migration

The setting section establishes the central puzzle: many students expect higher pecuniary value from destinations away from home, yet still prefer to study at home. In the conceptual framework in Section 3, this implies that a negative non-pecuniary value differential, $\Delta N_{i,c} < 0$, offsets the positive net pecuniary value differential, $\Delta V_{i,c} > 0$, for many students.

Using a discrete choice experiment and data on students' destination-specific beliefs, this section establishes three results about $\Delta N_{i,c}$, which I refer to as migration costs. First, they are large in stated preferences: in a discrete choice experiment that asks students

to trade off study location with labor-market returns, the average student would forgo a 129% increase in post-graduation earnings rather than study away from home. This amount increases with distance, from 39% for the closest option to 303% for the furthest option, and female students would forgo substantially more than male students. Second, they are anchored in systematic belief differentials between home and away destinations on specific non-pecuniary outcomes (above all safety, parental approval, and expected happiness), which together account for more than half of it. Third, students hold these beliefs in a sparse information environment: most have no source of college advice, few networks in destination cities, and limited knowledge of colleges away from home.

4.1 The choice experiment and elicitation of subjective beliefs

Choice experiment To estimate the non-pecuniary cost differential $\Delta N_{i,c}$, I use a discrete choice experiment (DCE) embedded in the baseline survey. Each student is presented with a sequence of pairwise choices between hypothetical college profiles in real locations, together with a third option, "No university". Each college is characterized by a major, a measure of educational quality, a destination within Maharashtra, a tuition fee, and monthly living expenses. Of these, the major, tuition, and living expenses are fixed. The major matches the student's preferred field, and tuition and living costs are held constant within bounds set by the student's previously stated budget. Students are also told that the two colleges are identical in all other aspects, and are instructed to assume they are admitted to both colleges and to choose the one they prefer, with not going to college at all as the only other option. Together, these features ensure that every college a student sees is feasible, both academically and financially, so admission chances and budget constraints do not drive students' choices.

Incentives The college profiles are hypothetical, but students' answers have consequences for them. The baseline survey doubles as the application to the mentorship program, and students and parents are told that their responses may be used to match students to mentors. Students are also told that a lottery decides who receives a mentor, so their answers affect only which mentor they receive. This gives students an incentive to answer truthfully, because misreporting cannot improve their chance of receiving a mentor and more accurate answers lead to a mentor better suited to their profile and needs (Kessler et al., 2019; Low, 2024). The same holds for the rest of the baseline survey, including the college rating exercise and the beliefs described below.

Identification Destination and educational quality vary randomly, and quality levels are set around the student's own previously elicited expectations. The exercise therefore asks students to trade off destination (*home* or somewhere *away*) against educational quality, holding all else constant. The profiles are not drawn from the actual college landscape: a student may face a top college at home against an average one in Pune even though no such college exists at home. In reality quality and destination are correlated (better colleges are disproportionately in the distant cities), so choices over realistic profiles could not separate a taste for quality from a preference for the destination itself; randomizing quality independently of destination breaks that correlation and identifies the disutility of migration.

Quality measures I implement this exercise in three quality-numeraire arms. The first, a natural numeraire, is expected earnings after graduation. However, this may not be the quality signal students actually observe when choosing colleges. I therefore also use job placement rates and admission cutoffs, which may be more readily available and guide students' choices. To reduce cognitive load, each student is randomly assigned to one arm and sees only one quality measure. Before the experiment, students watch a video that explains how to interpret the question and the profiles. Appendix F details the destination sets and the attribute levels in each arm, and shows the instructions and the screens students saw.

Estimation I specify utility as linearly additive in destination and the salary numeraire and estimate a conditional logit:

$$U_{ijt} = \beta_0 \text{NoUniversity}_j + \boldsymbol{\delta}'\mathbf{D}_j + \boldsymbol{\lambda}'\mathbf{S}_j + \varepsilon_{ijt},$$

where U_{ijt} is the latent utility student i assigns to alternative j in choice set t, $\mathbf{D}_j$ collects indicators for the destination of alternative j (the district capital, Pune, Mumbai, and Nagpur, or a single indicator for any away option), $\mathbf{S}_j$ the salary levels relative to the expected level (-25, $+25$, and $+50$ percent), and ε_{ijt} is type-I extreme value. Within a choice set the only varying attributes are destination and salary, so δ_d identifies the utility cost of destination d relative to home. Following Folke and Rickne (2022), the marginal utility of a one-percent salary increase is recovered from the level coefficients,

$$\hat{\beta}_{\%\text{Salary}} \approx \tfrac{1}{3}\left(-\tfrac{\hat{\lambda}_{-25}}{25} + \tfrac{\hat{\lambda}_{+25}}{25} + \tfrac{\hat{\lambda}_{+50}}{50}\right),$$

and the marginal rate of substitution, the salary increase that compensates a student for studying in d rather than at home, is $\text{MRS}_d \approx -\hat{\delta}_d \,/\, \hat{\beta}_{\%\text{Salary}}$. I use MRS_d as the estimate of migration costs defined in Section 3. Because tuition, living costs, and the major are held fixed, it measures only the non-pecuniary cost of studying in d rather than at home, expressed as a salary increase. The main estimates use the earnings arm, pooled with a midline re-administration of the same design to control students that adds the student's own district capital as a destination (Appendix F); the employment and cutoff arms, and migration costs in dollars, are reported in Appendix F.1.

College rating exercise I complement the choice experiment with a college rating exercise, also embedded in the baseline survey, and use it as a robustness check on the estimates of migration costs. To limit cognitive load, the choice experiment shows only the destination and one quality measure and holds everything else fixed, so students may still fill in details that are not shown with their own assumptions. In the rating exercise, students rate one college profile at a time from 0 to 10, by how likely they would be to accept admission if it were their only option. Each profile lists the destination together with fees, living costs, hostel information, the admission cutoff, the share placed in jobs, and average earnings, all drawn at random and independently of the student's own answers. The exercise therefore measures the same preference over destinations with more attributes shown, without anchoring to students' expectations, and with a different response format. Appendix G provides details.

Beliefs about destinations The baseline and midline surveys also ask how students picture student life at home and at two other randomly selected destinations along a set of non-pecuniary dimensions, which I group into three families. *Emotional well-being* is how happy they expect to be and how likely they are to struggle with their mental health. *Social ties* covers how often they would see family and friends back home, how many close friends they would make, whether people back home would think worse of them for leaving, and whether their parents would approve. *Practical and lived experience* covers how safe they would feel and how hard the move itself would be. Each item is elicited on a 0–10 scale, a probability, or a count, and I recode every item so that higher values are always better (safer, happier, more approving, and so on). I then standardize each by its home mean and home standard deviation, so a gap corresponds to how much worse students expect studying at a destination to be compared to home. Appendix H gives the exact question wording and response scales. Eliciting beliefs about every destination a student might consider is not feasible, and eliciting them only

for the destinations a student plans to study in would leave out the places the student is reluctant to consider, which is the margin this paper studies. The survey therefore assigns the other destinations, and Appendix E provides details.

4.2 Fact 1: Migration is costly

Table 1 reports the estimates of migration costs of leaving home (MRS_{Away}) and the coefficients behind them, overall and by gender, and Panels A and B of Figure 2 plot them by destination (Appendix Table F2). The away indicator is large, negative, and significant. The average student would forgo a 129% increase in post-graduation earnings (s.e. 26.3) rather than study away from home, and migration costs rise steeply with distance: 39% for the district capital, 93% for Pune, 196% for Mumbai, and 303% for Nagpur. When estimated with probability of employment as the numeraire, migration costs are equivalent to a 31.4 percentage-point increase in the probability of employment. The college rating exercise, which shows these attributes on each profile, implies migration costs of the same order (Appendix G).[16] The admission-cutoff arm points the same way, but I read it more cautiously. Taken at face value, migration costs are 253 program ranks, but students do not treat a higher cutoff as a desirable quality signal: they choose more-selective programs less. I therefore do not interpret the cutoff-arm willingness to pay as a quality tradeoff.[17]

Women would require a much larger increase in expected earnings than men to study away from home, 227% versus 72% (Panel B of Figure 2). This gap does not come from the utility cost of leaving home. The away coefficient, which measures this cost, is similar for men and women (-0.553 versus -0.614 in Table 1) and grows with distance at the same rate for both (Panel A of Appendix Table F2). An additional dollar of salary raises utility by about the same amount for men and women, so the dollar increase that compensates for leaving home is also about the same (column 4 of Appendix Table F3). Women expect to earn less than half as much as men (a median of \$90–180 versus \$270–360 a month), so one percent of their expected salary is worth less in utility

[16] Ratings respond to fees but only weakly to the earnings attribute, so I express migration costs in tuition. A college away from home is rated about a point lower, which corresponds to roughly \$920 a year in tuition, or about 86% of a year's median expected earnings, of the same order as the baseline DCE estimate of 126% (Appendix Table F3).

[17] Students' aversion to higher cutoffs could have several sources: they may not have fully internalized the instruction to assume they have scored high enough marks to be admitted to both profiles and may instead be reacting to the perceived difficulty of getting in; a higher cutoff may signal selectivity without a commensurate labor-market return; a more selective program may be seen as harder to study at; or applying to distant, more competitive programs may simply be more difficult. I cannot separate these. Only the first would threaten the identification of the disutility of migration, so as a robustness check I re-estimate the away coefficient in the pooled model, controlling for each student's perceived probability of admission at every destination. Even if students disregarded the assume-admission instruction entirely, the away coefficient falls from -0.59 to -0.38 (s.e. 0.05) in the same sample and remains highly significant. The same aversion to selectivity appears in the complementary college rating exercise, where the cutoff was drawn independently of the student's own expected score (Appendix G).

(the salary coefficient in Table 1), and the same dollar amount is a much larger share of their expected earnings. The college-rating exercise shows the same pattern.

These stated-preference estimates establish that migration costs are large, and that, as a share of expected earnings, they are much larger for women than for men. Large as they are, these costs are still smaller than the disutility of forgoing college altogether. In the pooled model, the coefficient on not attending university is two to three times the away coefficient, so students rank a college at home first, a college away second, and no university last.

Table 1: Migration costs estimated from the choice experiment

	All	Men	Women
Utility coefficients			
Away from home (δ_{Away})	-0.587***	-0.553***	-0.614***
	(0.054)	(0.085)	(0.070)
Ten percent of expected salary ($10 \times \beta_{\%\text{Salary}}$)	0.046***	0.077***	0.027***
	(0.008)	(0.013)	(0.010)
Migration costs: % salary increase needed to be indifferent between *home* and *away*			
$\text{MRS}_{\text{Away}} = -\delta_{\text{Away}} / \beta_{\%\text{Salary}}$	129***	72***	227**
	(26)	(16)	(93)
Observations	28,056	10,320	17,736
Students	981	365	616

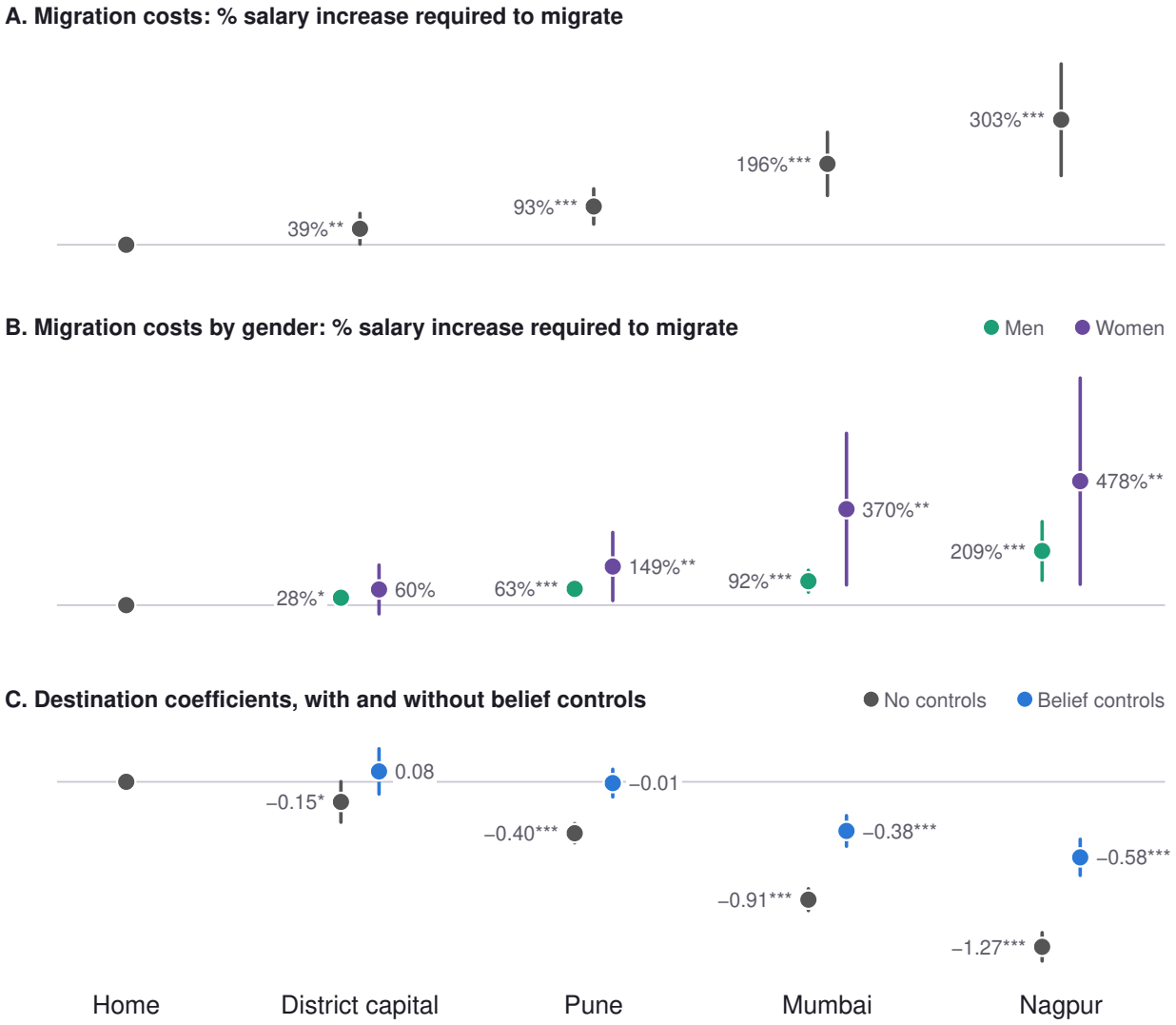
Notes: Conditional logit estimates from the choice experiment (Section 4.1). Standard errors clustered by student in parentheses, computed by the delta method for migration costs. The sample pools the baseline earnings arm with the midline control group. *Away from home* is an indicator for any destination other than home. *Ten percent of expected salary* is ten times the coefficient on a one percent increase in salary relative to the student's expected salary, averaged over the levels shown. *Migration costs* are the salary increase that compensates a student for studying away from home, the away coefficient divided by the coefficient on a one percent increase. Men and women are estimated separately. Observations are alternatives, three per choice including the option of not studying. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.3 Fact 2: Beliefs account for a large share of migration costs

Fact 1 shows that migration costs are large. Using the subjective beliefs about each destination elicited in the survey (Section 4.1), I next ask whether these costs reflect students' expectations about specific outcomes at different destinations.

Across all eight outcomes, students expect substantially worse experiences if they were to study away from home than if they studied at home (Figure 3, Table H2). The perceived decreases are largest for safety (-0.72 home standard deviations), parental approval (-0.62), and expected happiness (-0.53); the remaining decreases are smaller

Figure 2: Estimates of migration costs by destination, by gender, and with belief controls

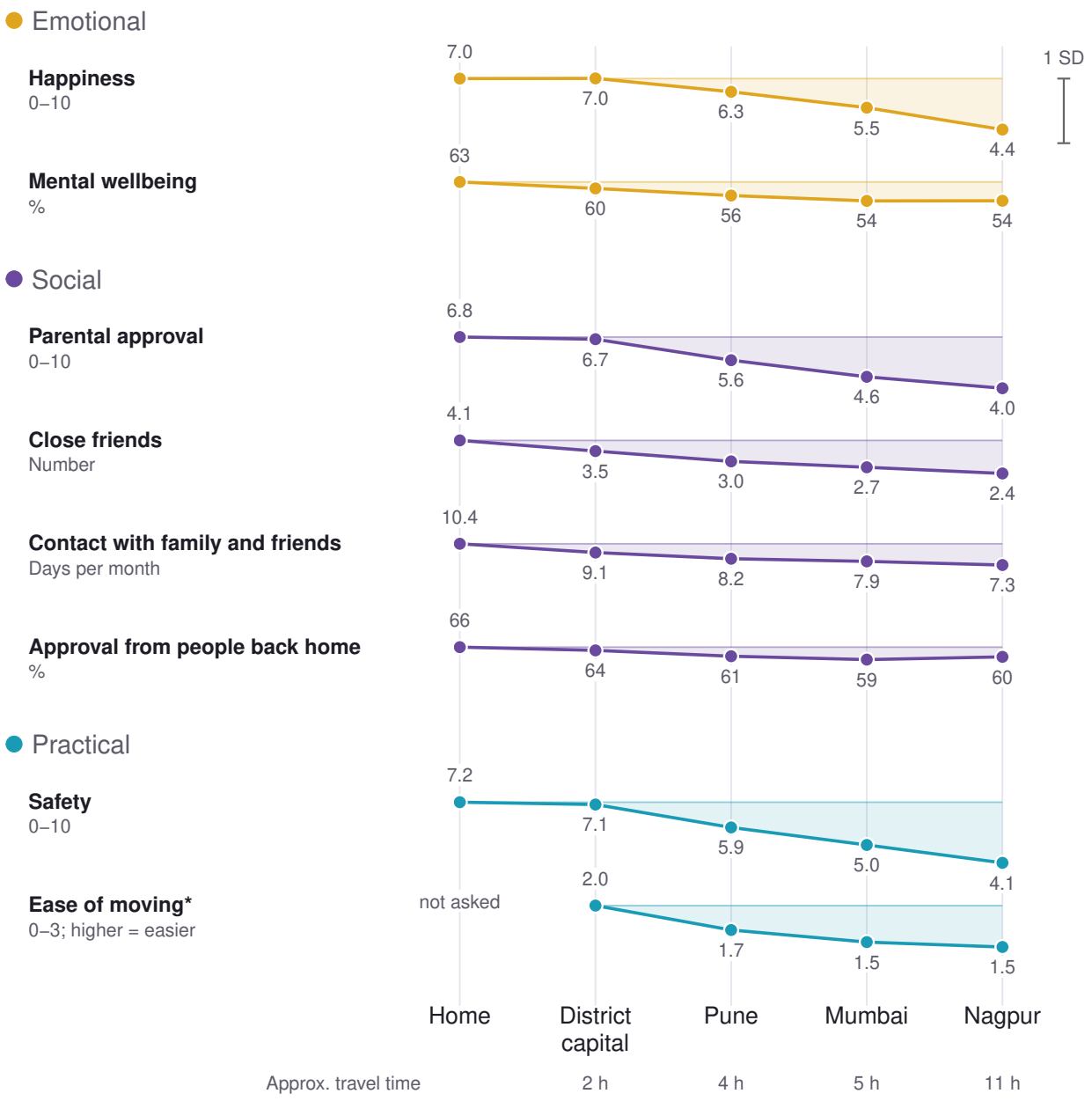


Notes: Estimates from the choice experiment (Section 4.1), with home as the reference. Panels A and B plot migration costs, in percent of expected salary. Panel C plots destination coefficients from the model with all three arms, without and with the eight beliefs as controls, on choices with complete belief data. Whiskers are 95% confidence intervals with standard errors clustered by student. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

but all significant at the 1% level. For women the decreases are larger for parental approval, safety, and reputation back home.

If these beliefs are part of the non-pecuniary value of destinations, controlling for them should shrink the migration costs estimated in Section 4.2. Panel C of Figure 2 re-estimates the destination coefficients with each student's own beliefs about each destination added as controls, using all three arms of the choice experiment for precision (Appendix F). The beliefs are jointly significant ($p < 0.001$) and absorb most of the disutility. In the model of Table 1, the migration costs of leaving home fall from 129% to

Figure 3: Beliefs about non-pecuniary outcomes by destination



Notes: The figure shows mean beliefs about each destination, ordered by distance from home. Travel times are approximate. Means pool the baseline survey and the midline control group and are adjusted for survey wave (Appendix E). Numbers are means in the units of each item, recoded so that higher values are better (Section 4.1). Vertical distances are in standard deviations of each item over all answers, and the bar shows one standard deviation. The shaded area is the decline from the first value in each row. *Ease of moving was not asked about home; its standard deviation uses away answers only.

42 % of post-graduation earnings, a reduction of 67.3 %. To see which beliefs account for this reduction, I decompose it following Gelbach (2016). I estimate a linear probability version of the pooled model. In this model, each belief's contribution to the fall in the away coefficient is its weight in choices times its average away–home gap, and the contributions add up to the total fall. Together the beliefs account for 61.3 % of the away coefficient, with the practical items accounting for the largest share (34.0 %), the emotional items for 13.3 %, and the social items for 14.1 % (Appendix Table F5).[18] Because a student's beliefs about a destination are correlated, the total contribution of beliefs is estimated more precisely than its split across families.

Beliefs account for a large share of the disutility of both men and women. In the college-rating exercise, beliefs account for nearly all of the away rating gap for both men and women. In the choice experiment, the share is larger for women in every arm and wave (74.1 % versus 41.6 % in the pooled model; Appendix Table F4). The disutility itself is similar for men and women, so the estimates suggest that the beliefs I elicit capture more of women's migration costs, and that men's rest more on factors these beliefs do not measure. The difference comes from the practical items, safety and the difficulty of moving (46.5 % versus 13.7 %; Appendix Table F5). These are the dimensions on which other studies find that women give up more than men, to stay close to home or to stay safe (Borker, 2021; Cheema et al., 2026; Folke and Rickne, 2022).

Migration costs are therefore large and systematic, and students' subjective beliefs about non-pecuniary outcomes at destinations account for a substantial share of them. As the next subsection shows, students have limited means of verifying these beliefs.

4.4 Fact 3: Students have limited information and few contacts

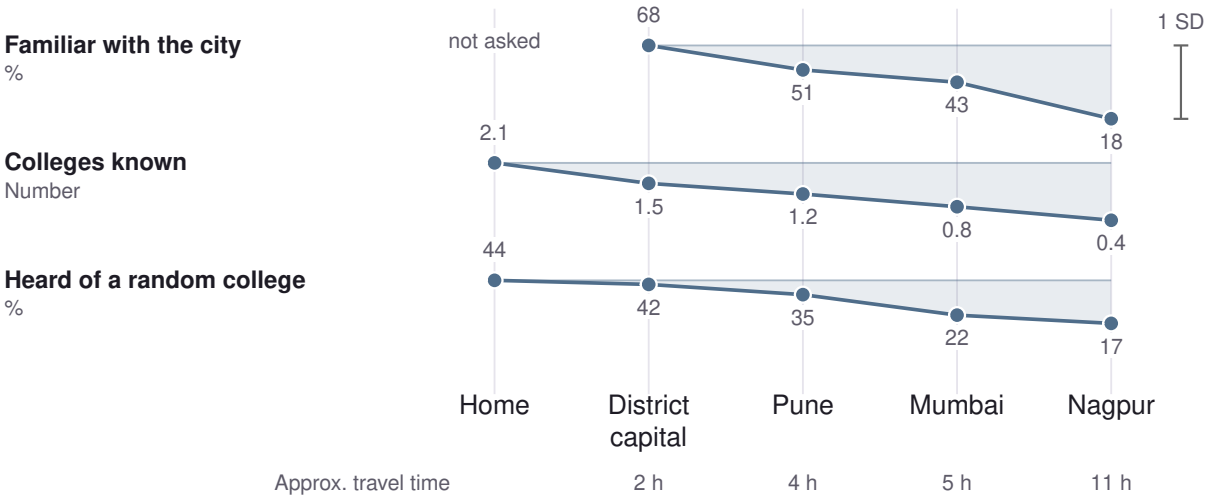
Facts 1 and 2 establish that migration costs are large and anchored in students' beliefs about non-pecuniary outcomes at the destination. Whether those beliefs are accurate, and whether they can be changed, depends on what students can observe about destinations and from whom. If students have little direct information about or exposure to the destinations, their beliefs may be inaccurate. This subsection documents that the broader information environment in which students form destination-specific beliefs is sparse: most students report low absolute levels of knowledge about both the college landscape and the destinations themselves, and few sources of information to learn about either.

[18] The share ranges from 36 % to 76 % across arms and waves (Appendix Table F4). Moving difficulty is asked only about away destinations and relative to home: the home option takes the value "not challenging at all," which 12 % of away answers also select, while 42 % describe the move as very or overwhelmingly challenging. The home value is definitional, so part of this item's contribution rests on the anchoring. Without the item, beliefs still account for 24 to 43 % of the away coefficient across arms and waves. In the college-rating exercise, beliefs account for the whole away rating gap, and for 81 % without moving difficulty (Appendix Tables F4 and G3).

At baseline, about 29 % of students report no source of college advice at all, and 30 % have not yet considered any specific colleges. Only 4 of the 52 schools in the sample employ a dedicated college counselor, so only about 12 % of students have institutional support inside the school. Even at home, students can on average name only 2.1 colleges in their preferred field, and when asked about a specific, randomly chosen college there, only 44 % have heard of it. Most students themselves report low confidence in navigating the admissions process (Appendix Figure H1).

Away from home, students know even less about colleges and studying away at other destinations (Figure 4). The count of colleges students know within their preferred field falls with distance, despite the fact that selective colleges are disproportionally located in these locations. Asked how well they know each of the cities the survey randomly assigns them, 60 % of the responses is "I don't know anything about it"; 41 % of students do not have any relatives or friends in any of the destinations asked. When asked which topics they would discuss with a current college student in a destination city, students choose cost of living (36 %), career prospects (35 %), safety (35 %), academics (34 %), living conditions (32 %), social life (23 %), family wishes (20 %) and adjusting to university life (13 %).

Figure 4: Knowledge of destinations and of colleges, by destination



Notes: The figure shows mean knowledge about each destination, ordered by distance from home. Travel times are approximate. Means pool the baseline survey and the midline control group and are adjusted for survey wave. Vertical distances are in standard deviations of each item over all answers, and the bar shows one standard deviation. The shaded area is the decline from the first value in each row. *Familiar with the city* is the share who know anything about studying and student life there. It was not asked about *home*; its standard deviation uses away answers only. *Colleges known* is the number of colleges in the student's preferred field that the student can name, top-coded at more than 10. *Heard of a random college* is the share who have heard of a college drawn at random from that destination.

Students who know someone in a destination know more about it and expect to be

better off there. This could be because contacts share information about the destination or lower the underlying migration costs, or because students with contacts differ from those without. Table 2 offers descriptive evidence that it is knowing someone in the particular destination that matters: Knowing someone in the destination closes about a third of the away–home gap in both non-pecuniary and knowledge indices, while knowing someone only in another destination does not (columns 3 and 6). The estimates may still reflect selection on destination-specific preferences.

Table 2: How beliefs and knowledge away from home differ by contacts in the destination

	Non-pecuniary index			Knowledge index		
	(1)	(2)	(3)	(4)	(5)	(6)
Away	-0.85***		-0.84***	-1.38***		-1.37***
	(0.08)		(0.08)	(0.06)		(0.06)
Away × knows someone in any away city	0.11		-0.11	0.29***		-0.08
	(0.10)		(0.11)	(0.09)		(0.11)
Away × knows someone in this city		0.28***	0.30***		0.52***	0.52***
		(0.06)	(0.05)		(0.07)	(0.06)
Destination fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Student-by-survey fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Home answers included	Yes	No	Yes	Yes	No	Yes
Students	1,682	1,682	1,682	1,682	1,682	1,682
Student–destination pairs	4,733	3,844	4,733	4,733	3,844	4,733

Notes: Each column reports an OLS regression of the index on the variables shown, with destination and student-by-survey fixed effects. Standard errors, clustered by student, in parentheses. The sample is the baseline survey (all students) and the midline survey (control group). A student knows someone in a destination if they have relatives, family friends, or friends living there. The non-pecuniary index combines expected happiness, safety, mental wellbeing, others not talking badly about them, days per month with family, number of close friends, and parental approval. The knowledge index combines familiarity with the city, the number of colleges in their preferred field the student can name there, and having heard of a randomly drawn college there. Familiarity is not asked about home, so home is coded as fully familiar. Both indices are inverse-covariance weighted and standardized on the away destinations of students who know no one in any away destination. Away is the average difference between away destinations and home for these students, weighted by the number of observations per destination. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Together, these three facts describe a setting in which students face large migration costs, anchored in beliefs about non-pecuniary outcomes they cannot easily verify. The randomized intervention described in Section 5 targets these large perceived migration costs by introducing students to peers who have already moved away from home to study. If the perceived migration costs are rooted in misperceptions, or fears, such a peer may offer new information that can correct students' beliefs. Alternatively, even if the beliefs are accurate on average, access to someone with destination experience may reduce the underlying cost of migration by making the destination easier to understand and navigate.

5 The mentorship experiment

The experiment asks whether contact with a peer who has moved away to study changes perceived migration costs and college choices. It does so through *Campus Trail*, a phone-based peer-mentorship program designed to expand students' access to migrant peers before college application deadlines. Eligible students are in their final year of secondary school and select into the study by signing up for the program when it is presented in their school. The sample is therefore selected on plans to continue to higher education, but is not conditional on wanting to study away from home (Section 2.2). Half of these students are randomly selected to be in the treatment group. Treated students are offered a spot in the mentorship program, and matched with a current university student in Pune or Mumbai who migrated for college and can discuss college choice, applications, and the practical experience of studying away from home. The program is therefore designed to target the perceived migration costs documented in Section 4.

Sampling I use a two-stage sampling process. First, I offer the program to every secondary school in the two study districts. Working from lists provided by the District Education Officers, the field team reaches each headmaster by phone to introduce the program and then visits the schools whose headmasters consent. Second, during a visit, the team presents Campus Trail to final-year students as a resource for preparing for college, describes the mentors as college students in Pune and Mumbai who moved there to study, leads a discussion to help students reflect on whether the program suits their goals, and invites interested students to sign up. Only students who sign up then complete the baseline survey, which also serves as their application to the program, on a tablet (Section 2.2). Appendix I describes how the program is presented.

Randomization The randomization sample consists of 1,682 students across 52 schools who meet pre-specified interest and attention criteria (Appendix B). I randomly assign half of the shortlisted applicants to *Campus Trail* and half to control, 826 and 856 students respectively. The two arms are balanced on baseline characteristics: no individual difference is significant, and a joint orthogonality test does not reject (Appendix Table B2).

Mentors and matching Mentors are current or recent university students in Pune and Mumbai who moved there from other parts of Maharashtra. Appendix D describes the applicant pool. They join as mentors after a short training on program rules, safeguarding, and the calling tools. Mentors' commitment is capped at three hours per mentee,

for which mentors receive a $2 contact bonus and a further $8 on completion (about $10 in total). Each treated student is matched to the best possible mentor for them. The mentor is always of the same gender and otherwise best fits the most preferred destinations and programs the student reports in the survey, as far as feasibility and mentors' availability allow. The pair then meets by phone. Calls start in September 2025, after the student's baseline survey, and most take place between October and December 2025, with a smaller second set in May and June 2026, after the entrance exams and as admissions begin (Appendix C). Conversations cover college and major choice and the practicalities of moving (housing, safety, finding one's way around an unfamiliar city), and draw on the mentor's own experience and knowledge of college admissions.

What mentors offer The mentor-applicant data suggest that these peers offer concrete, destination-specific knowledge as well as reassurance and encouragement; which correspond to the three channels outlined in Section 3 through which migration costs could decrease through contact with a migrant peer. First, mentors are prepared to offer concrete destination-specific information on budgets, accommodation, transport, safety, scholarships, and academic returns. Second, they could reduce effective migration costs by lowering search and coordination costs, for example by identifying hostels or paying-guest rooms, explaining procedures, suggesting routes, naming financial-support schemes, and connecting students to seniors or other contacts. Third, they could affect how students weight anticipated costs by acknowledging that moving may involve homesickness, loneliness, safety concerns, or parental resistance, while making compensating benefits salient, such as new friendships, broader peer networks, independence, exposure, and career opportunities. Appendix D documents these examples in detail.

Take-up Calls are conducted through a SurveyCTO integration that connects the pair while keeping mentors' contact details anonymous and logs each attempt's duration and outcome. I define take-up as at least fifteen minutes of conversation verified in these logs, following the definition in my pre-registration. The first conversation is scheduled by the field team, and enumerators attempted contact with every student assigned to the mentorship program. Of these students, 92% connected with their mentor at least once, and 31% went on to have a recorded conversation longer than fifteen minutes, averaging 41 minutes among them. These figures understate true engagement, because mentors sometimes called from personal phones or failed to record a call. Supplementing the logs with mentor-reported durations where possible implies that 45% of students had a call longer than fifteen minutes, with 40 minutes of engagement on average among them.

Among treated students who did not take up the intervention, enumerators logged an average of six scheduling attempts and mentors placed an average of five to six calls.

Who takes up Take-up follows students' own plans. A one-standard-deviation increase in a student's baseline likelihood of applying to colleges away from home raises take-up by about seven percentage points. Take-up also rises with academic preparation and self-efficacy, but not with household resources (Appendix Table B3).[19]

Students' reports of the calls Students who participated in conversations with their mentors reported positive impressions: among treated students who had spoken with their mentor by the midline survey, 91% agreed the mentor gave them useful information and advice, 90% felt encouraged after their sessions, and 85% said they trusted their mentor. Their conversations spanned the intended range of topics, most often academics (46%) and career prospects such as internships and placements (43%), but also the cost of living (24%), safety (22%), housing and daily logistics (20%), and navigating family wishes about studying away from home (19%).

Response and tracking 1,204 students (72%) completed the midline survey. The midline involved no tracking, so the midline sample is selected on school attendance on the survey day. At endline, if a student cannot be reached directly, a short proxy interview is administered with a parent or another family member. A quarter of the sample had not been reached by a pre-specified checkpoint. Following the pre-analysis plan, I randomly sampled ten schools with probability proportional to their number of unreached students, and the field team traced the unreached students in these schools through their classmates, recovering 92 of the 109 pursued. By the end of fieldwork, 1,143 students (68%) had completed a direct endline interview, rising to 1,418 (84%) once endline proxy interviews are included. With each recovered student weighted by the inverse probability that their school was drawn, the effective response rate rises from 75 percent at the checkpoint to 91 percent (Appendix Table J2). Response rates do not differ between the treatment and control groups at either wave (Appendix Table J1). Appendix Table J3 tests for selective and differential attrition, and Appendix J describes the endline tracking protocol and the attrition bounds. Parents of students with higher self-efficacy and

[19]These associations are individually and jointly significant in sample, but weak out of sample. A few baseline traits predict take-up (students who aspire to engineering or postgraduate study, or who are more willing to take risks, engage somewhat more), but across the full baseline covariate set take-up is essentially unpredictable out of sample (five-fold cross-validated $R^2 \approx 0$; random-forest out-of-bag $R^2 \approx 0.02$), as are total call length and mentors' engagement ratings, and engagement varies little across mentors (intra-class correlation ≈ 0.03).

stronger plans to apply away are more likely to complete the parent survey, but parent response does not differ across treatment arms (Appendix K).

6 Empirical strategy

This section describes the outcomes, the estimating equations, and the analysis samples, following the two pre-analysis plans filed before the midline and endline data were collected.

6.1 Endline

Primary outcome The primary outcome is an inverse-covariance-weighted index of standardized components (Anderson, 2008). The index combines six indicators measured at endline: applying to study, applying for education financing, enrolling in a program, applying to study away from home, migrating away from home, and migrating for education. Each indicator counts a realized action or, for students who have not yet acted, a planned application or move within the following two months. Each component is standardized to the control-group mean and standard deviation; the weights are the row sums of the inverted control-group covariance matrix; and the index is re-standardized to mean zero and unit standard deviation in the control group. It is computed for students with all six components observed. I report each outcome for any move away from home, the out-of-district definition is reported in Appendix Table L3.

Estimating equation For student i I estimate the intent-to-treat (ITT) effect by OLS:

$$Y_i = \alpha + \beta \text{CampusTrail}_i + \tilde{X}_i'\gamma + \left(\text{CampusTrail}_i \times \tilde{X}_i\right)'\delta + \varepsilon_i, \tag{1}$$

where CampusTrail_i indicates assignment to mentorship and β is the ITT. Following Lin (2013), the covariates are mean-centered ($\tilde{X}_i$) and interacted with treatment. X_i contains the baseline value of the outcome where available and additional controls chosen by post-double-selection lasso (Belloni et al., 2014) from a pre-specified candidate set (baseline and expected grades, assets, indices of agency, knowledge, and effort, time, risk, and competitiveness preferences, number of close friends, educational expectations, and baseline values of all outcomes), with missing baseline values set to zero and flagged by an indicator. I report heteroskedasticity-robust standard errors.

Beliefs about destinations At endline, students report their beliefs about studying at home and about one comparison destination, either their district capital or Pune. Students who live outside a district capital are randomly assigned one of the two; students who live in a district capital are asked about Pune. The student's realized or planned destination is itself an outcome of the intervention, and comparing beliefs about it across treatment arms would condition on a post-treatment outcome. The comparison destination is therefore assigned in advance. For each comparison destination, I stack each student's two observations and estimate

$$B_{id} = \alpha + \beta \text{CampusTrail}_i + \gamma \text{Away}_d + \delta \left(\text{CampusTrail}_i \times \text{Away}_d\right) + \mu_{g(i)} + \varepsilon_{id}, \quad (2)$$

where B_{id} is an index of student i's beliefs about destination d, home or the comparison destination, standardized on control students' beliefs about home; Away_d indicates the comparison destination; and $\mu_{g(i)}$ are gender-by-stream fixed effects, to which I add fixed effects for the student's district capital when it is the comparison destination. The coefficient of interest is δ, the effect of mentorship on beliefs about the comparison destination relative to beliefs about home. Observations are weighted by the inverse probability of being asked the endline belief module about the comparison destination; students who have moved or plan to move to another destination are asked about the comparison destination with probability 0.5, so their observations receive a weight of two. Standard errors are clustered by student. Expected salaries and monthly expenses are winsorized at the 99th percentile within treatment arm and destination, expressed in thousands of rupees, and transformed by the inverse hyperbolic sine.

Heterogeneous effects I report the pre-specified heterogeneity by gender, academic stream, and baseline contacts away from home. For each dimension, I interact assignment with an indicator G_i for one of the two groups:

$$Y_i = \alpha + \beta \text{CampusTrail}_i + \theta G_i + \delta \left(\text{CampusTrail}_i \times G_i\right) + \varepsilon_i, \quad (3)$$

where G_i indicates women, students in the commerce or arts stream, or students who at baseline knew someone (a relative, family friend, or friend) living in Pune, Mumbai, or Nagpur. The effect of mentorship is β in the reference group (men, science students, and students who knew no one away) and $\beta + \delta$ in the other group. Tables 5 and 6 and Appendix Tables M1 and M2 report the two effects in separate panels, each with the control mean of its group, and the p-value of δ in the last row. The index is the one in Table 3, standardized on all control students, so its control mean within a group need

not be zero. For beliefs, I estimate equation (2) separately for men and women. To test the difference, I interact every term of equation (2), including the fixed effects, with an indicator for being female. The last row of Table 7 reports the p-value of the coefficient on $\text{CampusTrail}_i \times \text{Away}_d \times \text{Female}_i$.

Analysis sample The analysis sample comprises all students with a completed endline interview, with the student or, when the student could not be reached, with a proxy. I also report estimates on direct student interviews only (Appendix Table L2).

Selective attrition Baseline characteristics do not jointly predict endline response ($p = 0.107$, Appendix Table J3). To limit concerns that the treatment effects are not representative of the study sample further, the field team traced unreached students in a random sample of schools through an intensive tracking procedure described in Section 5. Following the pre-analysis plan, I weight students recovered through intensive tracking by the inverse probability that their school was drawn (Appendix Table J2). Because the draw is random, the weighted sample also represents hard-to-reach students, who may differ from other students in ways the baseline survey does not capture.

Differential attrition Endline response rates are the same in both arms, but who responds differs (joint $p = 0.010$, Appendix Table J3). Treated students who intended to stay home at baseline are harder to reach than comparable control students, so treated respondents over-represent students inclined to migrate, which may bias the estimated effects on migration upwards. To check that the treatment effects are not driven by this, I report two sets of checks. First, following the pre-analysis plan, I bound the estimate by assigning worst-case values to students who were never reached or by shifting their outcomes by up to twice the observed difference between students reached in standard and in intensive tracking, in opposite directions across arms. Second, I assign each unreached student the outcome predicted for control students with the same baseline characteristics (Appendix Table J4).

6.2 Midline

Primary outcome The primary outcome at midline is students' willingness to migrate. For home and for each destination in the survey, students report the probability that they will apply to a college there, whether they plan to look for colleges there, and how they rate the destination on a scale from 0 to 10. I combine the three measures into an inverse-covariance-weighted index, standardized on control students' answers about home and

computed for students with all answers observed. Two further measures come from the choice experiment and the college-rating exercise of Section 4.1, which the midline survey repeats.

Estimating equation I stack each student's answers about home and about each destination away from home and estimate by OLS

$$W_{id} = \alpha + \beta \text{CampusTrail}_i + \gamma \text{Away}_d + \delta \left(\text{CampusTrail}_i \times \text{Away}_d\right) + \varepsilon_{id}, \qquad (4)$$

where W_{id} is one of the three measures, or the index, for student i and destination d, and Away_d indicates a destination away from home. The coefficient of interest is δ, the effect of mentorship on willingness to study away relative to home. For the choice experiment, I add the interactions of assignment with the away and opt-out indicators to the conditional logit of Section 4.1 with a single away indicator. For the college-rating exercise, I regress the rating of each college profile by OLS on its attributes, an indicator for a college away from home, assignment, and the interaction of assignment with the away indicator. In both, the coefficient on that interaction is the effect of mentorship on the value of studying away relative to home. Standard errors are clustered by student. I estimate each specification separately for men and women and test the difference by interacting every term with an indicator for being female (Appendix Table M4).

Analysis sample The analysis sample comprises all students who completed the midline survey. Baseline characteristics jointly predict midline response ($p = 0.001$). Women, students with higher past board marks, and students from larger households are more likely to be in school during the day of the midline data collection, hence to respond to the midline survey. Reassuringly, the baseline analogues of the primary outcomes do not predict response. Midline response rates are the same in both arms, and this selection does not differ by arm (joint $p = 0.362$) (Appendix Tables J1 and J3).

7 Effects of mentorship

7.1 Educational migration

Conversations with a migrant peer shift students' applications and realized moves. Table 3 reports intention-to-treat effects on the pre-specified migration index and its six components. Assignment to mentorship raises the index by 0.16 standard deviations. The index combines three indicators of continuing in higher education (applied for fur-

ther study, applied for financing, and enrolled) with three indicators of migration (applied to a college away from home, moved away from home, and moved away while enrolled). On the extensive margin, treated students are three percentage points more likely to have applied for further study and five points more likely to be enrolled, against control means of 91 and 79 percent. On the intensive margin, where they study, every estimate is positive but not statistically significant: four percentage points on applying to a college away from home and three on moving away. The estimate is stable across constructions of the index, from 0.13 to 0.20 standard deviations across the post-double-selection specification, restricting the sample to direct interviews with students only, including imcomplete interviews, and re-defining the outcome variables as realized outcomes only (Appendix Table L1); and it survives the attrition adjustments of the pre-analysis plan: 0.12 under intensive-tracking weights, and between 0.10 and 0.15 under adversarial sensitivity bounds (Appendix Table J4).

Table 3: Effects on educational migration

	Index of educational migration	Applied to study	Applied to education financing	Enrolled in a program	Applied to study away	Migrated away	Migrated for education
Assigned to mentorship	0.155***	0.031**	0.029	0.046**	0.040	0.027	0.018
	(0.054)	(0.014)	(0.027)	(0.021)	(0.026)	(0.022)	(0.021)
Control mean	0.000	0.906	0.521	0.788	0.335	0.216	0.181
Observations	1299	1391	1357	1383	1365	1403	1374

Notes: Each column reports the coefficient from an OLS regression of the outcome on an indicator for assignment to the mentorship program. Robust standard errors in parentheses. The sample is all students with a completed endline interview; for students who could not be interviewed, outcomes are reported by a parent or another family member. All outcomes except the index are indicators equal to one if the student has taken the action or plans to within two months. Away from home means outside the student's home sub-district. *Applied to study*: applied for further study. *Applied to education financing*: applied for a scholarship or loan. *Enrolled in a program*: enrolled in a degree college, vocational institute, or coaching program. *Applied to study away*: applied to study away from home. *Migrated away*: moved away from home. *Migrated for education*: migrated away and enrolled. *Index of educational migration*: inverse-covariance-weighted average of the six components (Anderson, 2008), for students with all six observed, standardized to mean zero and standard deviation one in the control group. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.2 Where the induced moves go

Consistent with the stated-preference evidence, the effect is concentrated on the shortest moves, where perceived migration costs are smallest. Table 4 breaks the applications and moves away from home down by destination. The additional applications increase only within the home district: five percentage points on a base of 18 percent, with no change for Pune and Mumbai or other districts. Moves follow the same pattern: moves within the home district increase by three percentage points, with no increases to the cities where the mentors study. These are also the destinations where migration costs are largest on average at baseline.

Table 4: Applications and moves by destination

	Applied to study				Migrated away		
	Within district		Out of district		Within district	Out of district	
	Home	Elsewhere	Pune or Mumbai	Elsewhere		Pune or Mumbai	Elsewhere
Assigned to mentorship	-0.017	0.049**	0.010	-0.012	0.031*	0.003	-0.008
	(0.026)	(0.022)	(0.015)	(0.015)	(0.017)	(0.013)	(0.012)
Control mean	0.617	0.179	0.084	0.096	0.096	0.057	0.059
Observations	1391	1391	1391	1391	1403	1403	1403

Notes: Each column reports the coefficient from an OLS regression of the outcome on an indicator for assignment to the mentorship program. Robust standard errors in parentheses. The sample is all students with a completed endline interview; for students who could not be interviewed, outcomes are reported by a parent or another family member. Outcomes are indicators equal to one if the student has applied to study in (*Applied to study*) or moved to (*Migrated away*) the destination in the column, or plans to within two months. *Home*: the student's home sub-district. *Elsewhere* within the district: the district capital or any other sub-district of the home district; moves within the district are to these places. *Elsewhere* out of the district: any district other than the home district, Pune, and Mumbai. A student can apply to several destinations. Seven students who moved or plan to move without naming a destination are in no column. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.3 Men and women

Men and women respond differently to the mentorship. Tables 5 and 6 split the estimates by gender. Men's migration index rises by 0.25 standard deviations, while the corresponding increase for women is a statistically insignificant 0.09 standard deviations.

Men's treatment effects are concentrated on the intensive margin. Mentorship does not change whether they continue to higher education, but it does change whether they migrate to do so. Men's applications away from home increase by ten percentage points, and their moves away from home by six, an imprecisely estimated increase. Their applications within the home district rise by ten percentage points and to Pune or Mumbai by five, and their moves within the home district by five.

Women show no effect on any migration outcome. Women's treatment effects are concentrated on the extensive margin: treated women are 2.7 percentage points more likely to have applied for higher education anywhere (from a base of 91 percent) and 8 percentage points more likely to have enrolled in higher education (from a base of 76 percent). When I estimate treatment effects on applications by destination, the patterns are suggestive of a displacement effect of women's applications. Treated women are less likely to apply to Pune or Mumbai and more likely to apply within the home district, although neither estimate is statistically significant. Midline data, collected before applications, suggest that this displacement was already present in the short run (Appendix Table M4). For women, the index of willingness to migrate and three of the five measures are negative, although only the decline in plans to look for colleges away from home is statistically significant. For men, every estimate is positive, and the effect in the choice experiment is statistically significant. The gender difference is statistically signifi-

cant only two of the measures, and men's decline in MRS_{Away} suggests that mentorship lowered migration costs for men but not for women. The pattern is the same when I exclude students who fail the pre-specified attention check.[20]

Table 5: Effects on educational migration, by gender

	Index of educational migration	Applied to study	Applied to education financing	Enrolled in a program	Applied to study away	Migrated away	Migrated for education
Panel A. Men							
Assigned to mentorship	0.245***	0.038	0.080*	-0.006	0.103**	0.063	0.013
	(0.089)	(0.024)	(0.044)	(0.032)	(0.043)	(0.039)	(0.038)
Control mean	0.066	0.895	0.459	0.840	0.398	0.269	0.249
Observations	494	532	526	528	520	539	524
Panel B. Women							
Assigned to mentorship	0.091	0.027	-0.000	0.078***	-0.003	0.001	0.018
	(0.067)	(0.018)	(0.035)	(0.027)	(0.031)	(0.026)	(0.025)
Control mean	-0.038	0.912	0.557	0.757	0.297	0.185	0.141
Observations	805	859	831	855	845	864	850
p-value, women = men	0.165	0.732	0.149	0.047	0.047	0.187	0.912

Notes: Each column reports estimates from an OLS regression of the outcome on an indicator for assignment to the mentorship program, an indicator for being female, and their interaction. Robust standard errors in parentheses. Panel A reports the coefficient on assignment (men). Panel B reports the sum of the coefficients on assignment and the interaction (women). The last row reports the *p*-value of the interaction coefficient. Control means are for the control group of each gender. The sample is all students with a completed endline interview; for students who could not be interviewed, outcomes are reported by a parent or another family member. All outcomes except the index are indicators equal to one if the student has taken the action or plans to within two months. Away from home means outside the student's home sub-district. *Applied to study*: applied for further study. *Applied to education financing*: applied for a scholarship or loan. *Enrolled in a program*: enrolled in a degree college, vocational institute, or coaching program. *Applied to study away*: applied to study away from home. *Migrated away*: moved away from home. *Migrated for education*: migrated away and enrolled. *Index of educational migration*: inverse-covariance-weighted average of the six components (Anderson, 2008), for students with all six observed, standardized to mean zero and standard deviation one in the control group. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[20]The pre-specified check flags the students who choose the same option position in all eight choice-experiment tasks, suggesting that students were straightlining, and corresponds to 14 percent of students at baseline.

Table 6: Applications and moves by destination, by gender

	Applied to study				Migrated away		
	Within district		Out of district		Within district	Out of district	
	Home	Elsewhere	Pune or Mumbai	Elsewhere		Pune or Mumbai	Elsewhere
Panel A. Men							
Assigned to mentorship	-0.060	0.098***	0.045*	-0.041	0.054*	0.006	-0.001
	(0.043)	(0.038)	(0.027)	(0.027)	(0.030)	(0.024)	(0.022)
Control mean	0.541	0.211	0.090	0.132	0.119	0.082	0.067
Observations	532	532	532	532	539	539	539
Panel B. Women							
Assigned to mentorship	0.013	0.017	-0.013	0.006	0.014	-0.001	-0.012
	(0.032)	(0.026)	(0.018)	(0.018)	(0.019)	(0.014)	(0.014)
Control mean	0.663	0.160	0.081	0.074	0.082	0.042	0.053
Observations	859	859	859	859	864	864	864
p-value, women = men	0.173	0.076	0.075	0.152	0.267	0.777	0.651

Notes: Each column reports estimates from an OLS regression of the outcome on an indicator for assignment to the mentorship program, an indicator for being female, and their interaction. Robust standard errors in parentheses. Panel A reports the coefficient on assignment (men). Panel B reports the sum of the coefficients on assignment and the interaction (women). The last row reports the *p*-value of the interaction coefficient. Control means are for the control group of each gender. The sample is all students with a completed endline interview; for students who could not be interviewed, outcomes are reported by a parent or another family member. Outcomes are indicators equal to one if the student has applied to study in (*Applied to study*) or moved to (*Migrated away*) the destination in the column, or plans to within two months. *Home*: the student's home sub-district. *Elsewhere* within the district: the district capital or any other sub-district of the home district; moves within the district are to these places. *Elsewhere* out of the district: any district other than the home district, Pune, and Mumbai. A student can apply to several destinations. Seven students who moved or plan to move without naming a destination are in no column. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.4 Beliefs about destinations

Mentorship could have changed behavior not only through changing students' expectations about non-pecuniary outcomes at destinations, but also through expectations about pecuniary outcomes, or through encouragement, support, and practical help. The estimates suggest that mentorship operates mainly through expectations. No other prespecified channel responds: knowledge of destinations, information about admissions, search effort, support, and traits are unchanged in both waves (Appendix Table M5 reports the traits; the others are not reported). I collect data on students' pecuniary and non-pecuniary expectations about studying at home and about one of the two most common destinations: their district capital, the nearest city and the destination of most treatment-induced applications and moves, or Pune (the closest of the major education hubs, and the one with the smallest migration costs of the out-of-district options I ask about at baseline). Table 7 reports the effects for all students and by gender, estimated from equation (2).

Men For men, I find that mentorship mostly changes how they perceive the pecuniary outcomes at different destinations. At endline, treated men expect monthly living costs

in the district capital to be lower relative to home, and returns after graduation to be higher in Pune. On average, both destinations look better to treated men on pecuniary grounds, relative to home, through lower costs or higher returns. On non-pecuniary outcomes, men's expectations about the district capital improve and those about Pune worsen, but neither treatment effect is statistically distinguishable from zero.

Women For women, none of the expectations I elicit improve. Their expectations of living costs at other destinations, relative to home, do not change, and expected pecuniary returns to studying away decrease. On non-pecuniary outcomes, there is no improvement, and, if anything, beliefs about life in Pune worsen relative to home. For women, the estimates suggest that talking to a mentor makes studying away look worse on both pecuniary and non-pecuniary dimensions. I read this as consistent with the large migration costs students perceive at baseline being broadly accurate rather than misperceived. If these beliefs were largely driven by misperceptions about what life might look like when studying away from home, contact with someone who has moved away to study could correct the beliefs, even if the resulting choices did not change. I find evidence to the contrary, which points to real migration costs, large enough to not be affected by a light-touch peer mentorship intervention, as a key barrier to educational mobility.

Suggestive mechanisms Data from mentors' call reports suggest a potential mechanism. Mentors report discussing money (scholarships, fees or living costs) with 32 percent of male mentees and 18 percent of female mentees, and scholarships with 18 and 8 percent. Some mentors also kept notes, available for about one in eight students. These notes provide anecdotal evidence in support of this mechanism: mentors explained state scholarship schemes, discussed living costs and hostel fees, and reminded students to get the caste and income certificates needed for fee concessions. This is consistent with the results for men, who expect lower living costs in the district capital (Table 7) and are more likely to apply for a scholarship or loan (Table 5). Mentors also talked about careers more often with men, which may help explain why men expect higher returns in Pune. With women, mentors talked more about safety, similar to Gallen and Wasserman (2026), who find that advice to women focuses more on work/life balance. In my setting, mentors and students were matched by gender, and I cannot separate whether this pattern is a result of male mentees asking more advice about careers and costs, or male mentors offering information about it more.

Table 7: Effects on beliefs about destinations (endline)

	Within district			Out of district		
	Non-pecuniary	Returns	Costs	Non-pecuniary	Returns	Costs
Panel A. All students						
Assigned to mentorship × away	0.055	-0.117	-0.144	-0.135	-0.003	-0.063
	(0.119)	(0.097)	(0.092)	(0.117)	(0.100)	(0.090)
Control mean, away − home	-0.804	0.753	0.887	-1.062	0.804	1.204
Observations	931	931	891	975	975	929
Panel B. Men						
Assigned to mentorship × away	0.118	-0.139	-0.419***	-0.105	0.309*	-0.015
	(0.176)	(0.150)	(0.132)	(0.200)	(0.188)	(0.130)
Control mean, away − home	-0.850	0.666	0.903	-1.018	0.754	1.111
Observations	409	409	393	346	346	332
Panel C. Women						
Assigned to mentorship × away	0.006	-0.081	0.101	-0.158	-0.196*	-0.084
	(0.165)	(0.126)	(0.127)	(0.143)	(0.113)	(0.120)
Control mean, away − home	-0.771	0.816	0.876	-1.086	0.832	1.254
Observations	522	522	498	629	629	597
p-value, women = men	0.642	0.771	0.004	0.828	0.021	0.697

Notes: Each column reports estimates from an OLS regression of a belief index on an indicator for assignment to the mentorship program, an indicator for the destination (away), and their interaction, with gender-by-stream fixed effects. Each student contributes two observations: beliefs about home and beliefs about the destination in the column: the student's district capital (within district) or Pune (out of district). The within-district columns also include fixed effects for the student's district capital. Observations are weighted with the sampling weights described in Section 6. Standard errors clustered by student in parentheses. The first row reports the coefficient on the interaction. *Control mean, away − home*: the coefficient on the away indicator, the control group's belief about the destination relative to home. Panels B and C estimate the regression separately for men and for women. The last row reports the p-value of the coefficient on the triple interaction with an indicator for being female, from a regression in which all terms and fixed effects are interacted with gender. Indices are inverse-covariance-weighted averages (Anderson, 2008), standardized on control students' beliefs about home. *Non-pecuniary*: anticipated happiness, safety, mental wellbeing, whether people back home would think worse of the student, days per month with family, number of close friends, and parental approval, each signed so that higher is better. *Returns*: expected monthly salary after graduating and probability of employment. *Costs*: expected monthly expenses, living costs plus one twelfth of yearly tuition; higher means more expensive. Salary and expenses are winsorized at the 99th percentile within treatment arm and destination and enter as the inverse hyperbolic sine of thousands of rupees. *Observations*: student-destination observations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.5 The role of parents

Parents play an important role in young people's education decisions. Parents and children often make these choices jointly (Giustinelli, 2016), their preferences can conflict (Bursztyn and Coffman, 2012; Ashraf et al., 2020), and mothers' expectations can predict their daughters' schooling choices when the daughters' own expectations do not (Attanasio and Kaufmann, 2014). In my sample, too, most students do not decide alone. At baseline, I ask students who decides whether they continue in school after grade 12, whether they work after their studies, what type of work they do, which college they attend, which city they study in, and what they study. For the city, 26 percent of students

decide by themselves, 56 percent decide together with their parents, and 18 percent say that their parents decide for them. These decisions are gendered. Women are more likely than men to say that their parents decide which city they will study in (21 versus 15 percent), which college they will attend (19 versus 15 percent), and whether they continue after grade 12 (18 versus 14 percent). This is in line with evidence that gender norms shape families' decisions in India (Jayachandran, 2015).

Given how central parents are to their children's education decisions, I also survey one parent or guardian per student at baseline (Section 2.2). Parents are even more averse to migration than their children. In the same choice experiment, the average parent would forgo about a 450 percent increase in the child's post-graduation earnings to avoid their child moving outside of their home district to study, compared with about 140 percent for students, and parents expect a larger loss than students on every non-pecuniary dimension I elicit (Appendix K). Even though the parents are informed about the mentorship program, the program does not target and engage parents directly.[21] One may therefore worry that a program aimed at students can do little when parents oppose migration, and that women's null effects on migration reflect constraints set by their parents.

If parents are the main constraint on students' decisions to migrate, a program that reaches only the students should have larger effects on students with greater decision-making power over their own education, assuming that these students can act on their wishes without parental constraints. To test this, I combine the six answers into a standardized index of decision-making power and compare students above and below its median (Appendix Table M3). Among men, the effects of mentorship are the same in both groups. Men with less decision-making power respond as much as men with more, so the effects of the program likely reach the parents of the former group.

Women's results speak to the second concern. If women wanted to study away after talking to a mentor but their parents kept them at home, women with greater decision-making power should respond more than women with less. They do not. This is more consistent with women choosing not to migrate even when the decision is theirs.

8 Conclusion

This paper studies one of the first and most consequential migration decisions young people take, at a critical point in life: whether and where to study for college. I follow 1,682 high-school students in two districts of Maharashtra as they finish school and

[21] Anecdotal reports from mentors suggest that some mentors had to speak with a student's parents, and get their approval, before they could speak with the student. The program may therefore have reached some parents indirectly.

apply to and enroll in higher education. In an incentivized choice experiment, I measure perceived migration costs and the beliefs behind them. In a randomized controlled trial, I evaluate a mentorship program that connects these students with college students who have left home to study.

Among these students, perceived migration costs are large. The average student would forgo a 129% increase in expected post-graduation earnings to avoid studying away from home, and beliefs about the difficulty of the move, safety, parental approval, and happiness account for a large share of these costs. These perceived migration costs far exceed the expected returns from studying elsewhere, and students plan to study at home. Migration costs are larger for women, who expect larger declines in safety, parental approval, and reputation back home if they were to move. Students have few sources of information to verify these beliefs against. Those who know someone at a destination hold more favorable beliefs about living and studying at that destination, suggesting that contacts can decrease perceived migration costs, but many have no such contact.

A light-touch intervention consisting of conversations with a college student who is an educational migrant changes choices, but only at the margin where the perceived cost was small to begin with. Men respond by migrating to study, while women respond by applying to and enrolling in higher education at home, with their applications shifting, if anything, from further away destinations toward nearer ones. The intervention does not change migration costs for the distant cities, where the pecuniary gains concentrate, and beliefs about life there do not improve and.

These findings have implications on how to widen access to educational opportunity. Bringing colleges closer to students may raise enrollment in higher education locally, but if migration costs still bind, graduates will largely enter the labor market where they studied (Conzelmann et al., 2025). In that case, moving students to education may instead be an alternative. Ghose (2026) estimates that in India, lowering the cost of moving for education would raise welfare by more than lowering the cost of moving for work. However, the results of the two experiments indicate that migration costs are a major obstacle to these moves. At the margin where migration costs are relatively low, information on living costs or financial support may help students migrate to opportunity. Moving far, to places where migration costs are large, is different; and a light-touch intervention does not change how students perceive it. If beliefs about life in the distant cities are broadly correct, and the fact that they do not improve on contact with someone who has moved away to study suggests they are, encouraging educational migration at scale might require other policy tools that decrease migration costs by making cities

safer, easier to navigate and providing potential migrants intensive support that continue after they arrive.

References

Alba-Vivar, F. M. (2025). Opportunity bound: Transport and access to college in a megacity. Working Paper 128, Department of Economics, Wake Forest University.

Alfonsi, L., M. Namubiru, and S. Spaziani (2024). Meet your future: Experimental evidence on the labor market effects of mentors. Working Paper 87, G^{2}LM|LIC.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Ashraf, N., N. Bau, C. Low, and K. McGinn (2020). Negotiating a better future: How interpersonal skills facilitate intergenerational investment. *Quarterly Journal of Economics 135*(2), 1095–1151.

Attanasio, O. P. and K. M. Kaufmann (2014). Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics 109*, 203–216.

Bah, T. L., C. Batista, F. Gubert, and D. McKenzie (2023). Can information and alternatives to irregular migration reduce "backway" migration from The Gambia? *Journal of Development Economics 165*, 103153.

Baláž, V., A. M. Williams, and E. Fifeková (2016). Migration decision making as complex choice: Eliciting decision weights under conditions of imperfect and complex information through experimental methods. *Population, Space and Place 22*(1), 36–53.

Barnett-Howell, Z., T. Baseler, T. Ginn, and S. Gordeev (2026). Reaching the novice or nudging the expert? Information delivery and the returns to migration. Working paper.

Baseler, T. (2023). Hidden income and the perceived returns to migration. *American Economic Journal: Applied Economics 15*(4), 321–352.

Batista, C., D. M. Costa, P. Freitas, G. Lima, and A. B. Reis (2025). What matters for the decision to study abroad? A lab-in-the-field experiment in Cape Verde. *Journal of Development Economics 173*, 103401.

Batista, C. and D. McKenzie (2023). Testing classic theories of migration in the lab. *Journal of International Economics 145*, 103826.

Battiston, D. (2018). The persistent effects of brief interactions: Evidence from immigrant ships. Working paper, MPRA Paper No. 97151.

Battiston, G., L. Corno, and E. La Ferrara (2024). Informing risky migration: Evidence from a field experiment in Guinea. RFBerlin Discussion Paper 2434, ROCKWOOL Foundation Berlin.

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in India. *Science 335*(6068), 582–586.

Belloni, A., V. Chernozhukov, and C. Hansen (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives 28*(2), 29–50.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012). The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *Quarterly Journal of Economics 127*(3), 1205–1242.

Bhatiya, A. Y., B. Chakravorty, C. Imbert, and R. Rathelot (2026). How do beliefs and preferences over jobs affect enrollment in vocational training: Experimental evidence from India. *European Economic Review 184*, 105238.

Blumenstock, J. E., G. Chi, and X. Tan (2025). Migration and the value of social networks. *Review of Economic Studies 92*(1), 97–128.

Boneva, T. and C. Rauh (2017). Socio-economic gaps in university enrollment: The role of perceived pecuniary and non-pecuniary returns. CESifo Working Paper 6756, CESifo.

Borker, G. (2021). Safety first: Perceived risk of street harassment and educational choices of women. Policy Research Working Paper 9731, World Bank.

Breda, T., J. Grenet, M. Monnet, and C. Van Effenterre (2023). How effective are female role models in steering girls towards STEM? Evidence from French high schools. *Economic Journal 133*(653), 1773–1809.

Bryan, G., S. Chowdhury, and A. M. Mobarak (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica 82*(5), 1671–1748.

Bryan, G. and M. Morten (2019). The aggregate productivity effects of internal migration: Evidence from Indonesia. *Journal of Political Economy 127*(5), 2229–2268.

Bursztyn, L. and L. C. Coffman (2012). The schooling decision: Family preferences, intergenerational conflict, and moral hazard in the Brazilian favelas. *Journal of Political Economy 120*(3), 359–397.

Caner, A., M. Derebasoglu, and C. Okten (2024). Attainment and gender equality in higher education: Evidence from a large-scale expansion. *Journal of Human Capital 18*(3), 469–530.

Carrell, S. E. and B. Sacerdote (2017). Why do college-going interventions work? *American Economic Journal: Applied Economics 9*(3), 124–151.

Carrington, W. J., E. Detragiache, and T. Vishwanath (1996). Migration with endogenous moving costs. *American Economic Review 86*(4), 909–930.

Castleman, B. L. and L. C. Page (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior & Organization 115*, 144–160.

Chandrasekhar, S. and A. Sharma (2015). Urbanization and spatial patterns of internal migration in India. *Spatial Demography 3*(2), 63–89.

Chauvin, J. P., E. Glaeser, Y. Ma, and K. Tobio (2017). What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics 98*, 17–49.

Cheema, A., A. I. Khwaja, M. F. Naseer, and J. N. Shapiro (2026). Glass walls: Experimental evidence on constraints faced by women in accessing valuable skilling opportunities. *Journal of Political Economy 134*(6), 1775–1845.

Clayton, K., Y. Horiuchi, A. R. Kaufman, G. King, and M. Komisarchik (2026). Correcting measurement error bias in conjoint survey experiments. *American Journal of Political Science*. Forthcoming.

Clerici, C., I. Bianchi, and D. Biesalski (2024). Information, expectations and preferences: Occupational choices of young adults in Uganda. Working paper, Stockholm School of Economics.

Conzelmann, J. G., S. W. Hemelt, B. J. Hershbein, S. Martin, A. Simon, and K. M. Stange (2025). Grads on the go: Measuring college-specific labor markets for graduates. *Journal of Policy Analysis and Management 44*(3), 741–763.

Coppock, A., A. S. Gerber, D. P. Green, and H. L. Kern (2017). Combining double sampling and bounds to address nonignorable missing outcomes in randomized experiments. *Political Analysis 25*(2), 188–206.

Dean, J. T., G. Kreindler, and O. Mbonu (2026). The effect of exposure: Evidence from spatial choices in Nairobi. NBER Working Paper 34799, National Bureau of Economic Research.

Delavande, A. and B. Zafar (2019). University choice: The role of expected earnings, nonpecuniary outcomes, and financial constraints. *Journal of Political Economy 127*(5), 2343–2393.

Demirel, M., X. Hou, and Y. Kanoria (2026). Distance, gender, and access to elite higher education. Working paper.

Elsayed, A. and A. Shirshikova (2023). The women-empowering effect of higher education. *Journal of Development Economics 163*, 103101.

Fabre, A. (2024). The geography of higher education and spatial inequalities. Working paper.

Farré, L. and F. Ortega (2024). Geographic mobility of college students and the gender gap in academic aspirations. *Labour Economics 90*, 102550.

Folke, O. and J. Rickne (2022). Sexual harassment and gender inequality in the labor market. *Quarterly Journal of Economics 137*(4), 2163–2212.

Frohnweiler, S., B. Beber, and C. Ebert (2024). Information frictions, belief updating and internal migration: Evidence from Ghana and Uganda. *Journal of Development Economics 171*, 103311.

Fu, C., J. Guo, A. J. Smith, and A. T. Sorensen (2022). Students' heterogeneous preferences and the uneven spatial distribution of colleges. *Journal of Monetary Economics 129*, 49–64.

Gallen, Y. and M. Wasserman (2026). Informed choices: Gender gaps in career advice. NBER Working Paper 35750, National Bureau of Economic Research.

Garlick, R., E. Field, and K. Vyborny (2025). Women's mobility and labor supply: Experimental evidence from Pakistan. IZA Discussion Paper 17883, Institute of Labor Economics (IZA).

Gelbach, J. B. (2016). When do covariates matter? And which ones, and how much? *Journal of Labor Economics 34*(2), 509–543.

Ghose, D. (2026). Trade, internal migration, and human capital: Who gains from India's IT boom? *Review of Economics and Statistics*. Forthcoming.

Giustinelli, P. (2016). Group decision making with uncertain outcomes: Unpacking child–parent choice of the high school track. *International Economic Review 57*(2), 573–602.

Giustinelli, P. (2023). Expectations in education. In R. Bachmann, G. Topa, and W. van der Klaauw (Eds.), *Handbook of Economic Expectations*, pp. 193–224. Academic Press.

Gollin, D., D. Lagakos, and M. E. Waugh (2014). The agricultural productivity gap. *Quarterly Journal of Economics 129*(2), 939–993.

Green, A. (2024). Networks and geographic mobility: Evidence from World War II navy ships. Job Market Paper, Princeton University. Available at `https://allie-e-green.github.io/public/WW2_AG_Migration_jmp.pdf`.

Han, Y., D. Kübler, Y. Liu, X. Zheng, and Y. Zong (2025). The gender gap in gender-blind university admissions. Discussion Paper 62, Berlin School of Economics.

Hole, A. R. (2015). DCREATE: Stata module to create efficient designs for discrete choice experiments. Statistical Software Components S458059, Boston College Department of Economics. Revised 2017.

Hoxby, C. M. and S. Turner (2013). Expanding college opportunities for high-achieving,

low income students. Discussion Paper 12-014, Stanford Institute for Economic Policy Research.

Imbert, C. and J. Papp (2020). Costs and benefits of rural-urban migration: Evidence from India. *Journal of Development Economics 146*, 102473.

Ishimaru, S. (2025). Geographic mobility of youth and spatial gaps in local college and labor market opportunities. *Journal of Labor Economics 43*(4), 1251–1294.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics 7*(1), 63–88.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics 125*(2), 515–548.

Kapoor, R. and S. Gade (2025). Together in search: Experimental evidence from coordinating travel among women job-seekers in urban India. Working Paper 90, G^2LM|LIC.

Katzkowicz, N., V. Lavy, M. Querejeta, and T. Rosá (2026). Human capital consequences of expanding the supply of higher education institutions in a developing country. *Journal of Public Economics 260*, 105706.

Kennan, J. and J. R. Walker (2011). The effect of expected income on individual migration decisions. *Econometrica 79*(1), 211–251.

Kessler, J. B., C. Low, and C. D. Sullivan (2019). Incentivized resume rating: Eliciting employer preferences without deception. *American Economic Review 109*(11), 3713–3744.

Koenen, M. and D. Johnston (2024). Social networks and residential choice: Micro evidence and equilibrium implications. Working paper.

Koşar, G., T. Ransom, and W. Van der Klaauw (2022). Understanding migration aversion using elicited counterfactual choice probabilities. *Journal of Econometrics 231*(1), 123–147.

Lagakos, D. (2020). Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic Perspectives 34*(3), 174–192.

Lagakos, D., A. M. Mobarak, and M. E. Waugh (2023). The welfare effects of encouraging rural-urban migration. *Econometrica 91*(3), 803–837.

Le Barbanchon, T., R. Rathelot, and A. Roulet (2021). Gender differences in job search: Trading off commute against wage. *Quarterly Journal of Economics 136*(1), 381–426.

Lin, W. (2013). Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique. *Annals of Applied Statistics 7*(1), 295–318.

Low, C. (2024). Pricing the biological clock: The marriage market costs of aging to women. *Journal of Labor Economics 42*(2), 395–426.

Mangal, K. (2024). How much are government jobs in developing countries worth? *World Bank Economic Review 38*(3), 443–465.

Manski, C. F. (2004). Measuring expectations. *Econometrica 72*(5), 1329–1376.

McKenzie, D. (2024). Fears and tears: Should more people be moving within and from developing countries, and what stops this movement? *World Bank Research Observer 39*(1), 75–96.

McKenzie, D., J. Gibson, and S. Stillman (2013). A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad? *Journal of Development Economics 102*, 116–127.

Mesplé-Somps, S. and B. Nilsson (2023). Role models, aspirations and desire to migrate. *Journal of Economic Behavior & Organization 212*, 819–839.

Ministry of Human Resource Development (2013). All India Survey on Higher Education 2011–12. Technical report, Department of Higher Education, Government of India, New Delhi.

Munshi, K. (2020). Social networks and migration. *Annual Review of Economics 12*, 503–524.

Munshi, K. and M. Rosenzweig (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review 106*(1), 46–98.

Muralidharan, K. and N. Prakash (2017). Cycling to school: Increasing secondary school enrollment for girls in India. *American Economic Journal: Applied Economics 9*(3), 321–350.

Patnaik, A. and K. Mahalingam (2026). A district-level dataset of building volume and nighttime lights. The LEAP Blog, 22 October.

Riley, E. (2024). Role models in movies: The impact of Queen of Katwe on students' educational attainment. *Review of Economics and Statistics 106*(2), 334–351.

Sahai, H. and M. Bailey (2022). Social networks and spatial mobility: Evidence from Facebook in India. arXiv preprint 2203.05595.

Sarvimäki, M., R. Uusitalo, and M. Jäntti (2022). Habit formation and the misallocation of labor: Evidence from forced migrations. *Journal of the European Economic Association 20*(6), 2497–2539.

Sekhri, S. (2020). Prestige matters: Wage premium and value addition in elite colleges. *American Economic Journal: Applied Economics 12*(3), 207–225.

Sekhri, S., M. A. Hossain, and P. Khosla (2024). Access to colleges, human capital, and empowerment of women. *Journal of Human Resources 59*(2), 502–544.

Shrestha, M. (2020). Get rich or die tryin': Perceived earnings, perceived mortality rates,

and migration decisions of potential work migrants from Nepal. *World Bank Economic Review 34*(1), 1–27.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy 70*(5), 80–93.

Tjaden, J. and F. A. Dunsch (2021). The effect of peer-to-peer risk information on potential migrants: Evidence from a randomized controlled trial in Senegal. *World Development 145*, 105488.

Wiswall, M. and B. Zafar (2015). Determinants of college major choice: Identification using an information experiment. *Review of Economic Studies 82*(2), 791–824.

A Appendix: Higher education and earnings in Maharashtra

This appendix documents the concentration of higher-quality higher-education supply in Pune and the Mumbai region, where the mentors study, relative to the study districts (Ahilyanagar and Sangli), and provides the institutional context referenced in Section 2. The figures show where colleges are located, how many of their seats a student in the sample could obtain, how many other candidates compete for those seats, and how selective the programs are. The last table compares incomes in Pune and the Mumbai region with the rest of the state.

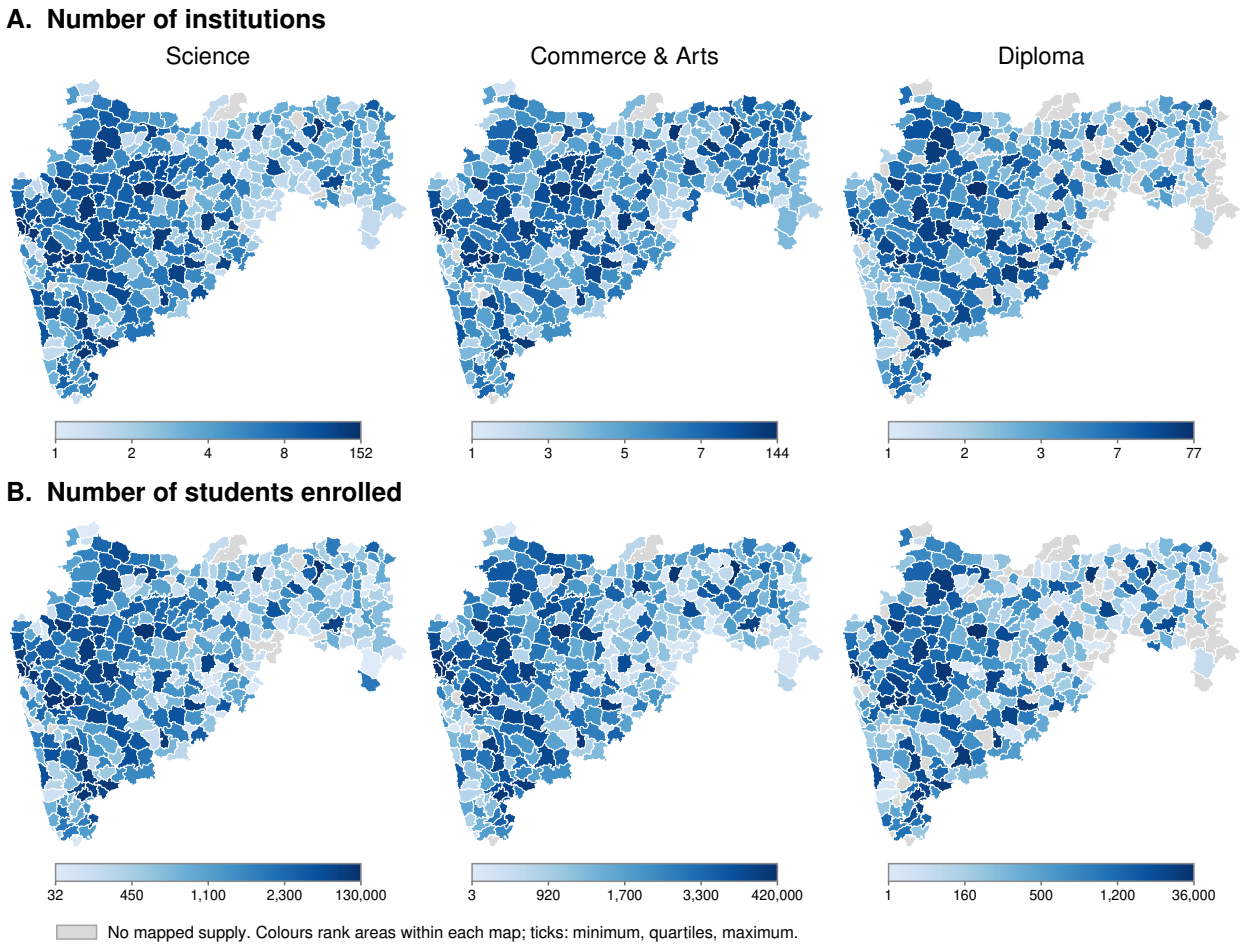
A.1 Supply and access

Supply across Maharashtra. Figure A1 maps higher-education supply across Maharashtra's sub-districts, separately for science programs, commerce-and-arts programs and diplomas. Supply varies widely across sub-districts. These numbers do not show which seats a given student could obtain.

Centralized admissions and quotas. Engineering, pharmacy, and business seats in Maharashtra are allocated primarily through the Maharashtra Common Entrance Test (MHT-CET), a state-administered exam taken after twelfth grade (business programs use a separate state exam, the MAH-BBA/BMS/BBM CET). Allocation to specific programs is centralized: candidates submit ranked program preferences and are assigned to seats based on their entrance-exam score, subject to category- and state-level quotas that reserve a substantial share of seats for in-state applicants, and to a home-university quota that reserves most seats in non-autonomous colleges for applicants from the college's university region. Medical and elite-engineering seats are allocated through separate national exams (NEET and JEE, respectively). The figures below use 2025–26 admissions data from the State Common Entrance Test Cell. Cutoffs and quotas are observed only for programs filled through the centralized system, so the figures that follow cover only these programs. They leave out general degree colleges (B.A., B.Com. and B.Sc.), which admit students directly.

Accessible seats. Figure A2 counts the seats that students in the sample could obtain, given admission cutoffs and quotas, if they scored at the median of each entrance exam. On average, such a student can reach about 115 engineering seats at home, 610 within their district and 7,000 within Maharashtra; for business programs the counts are about 55, 410 and 5,900.

Figure A1: Higher-education supply across Maharashtra, all colleges, by program group

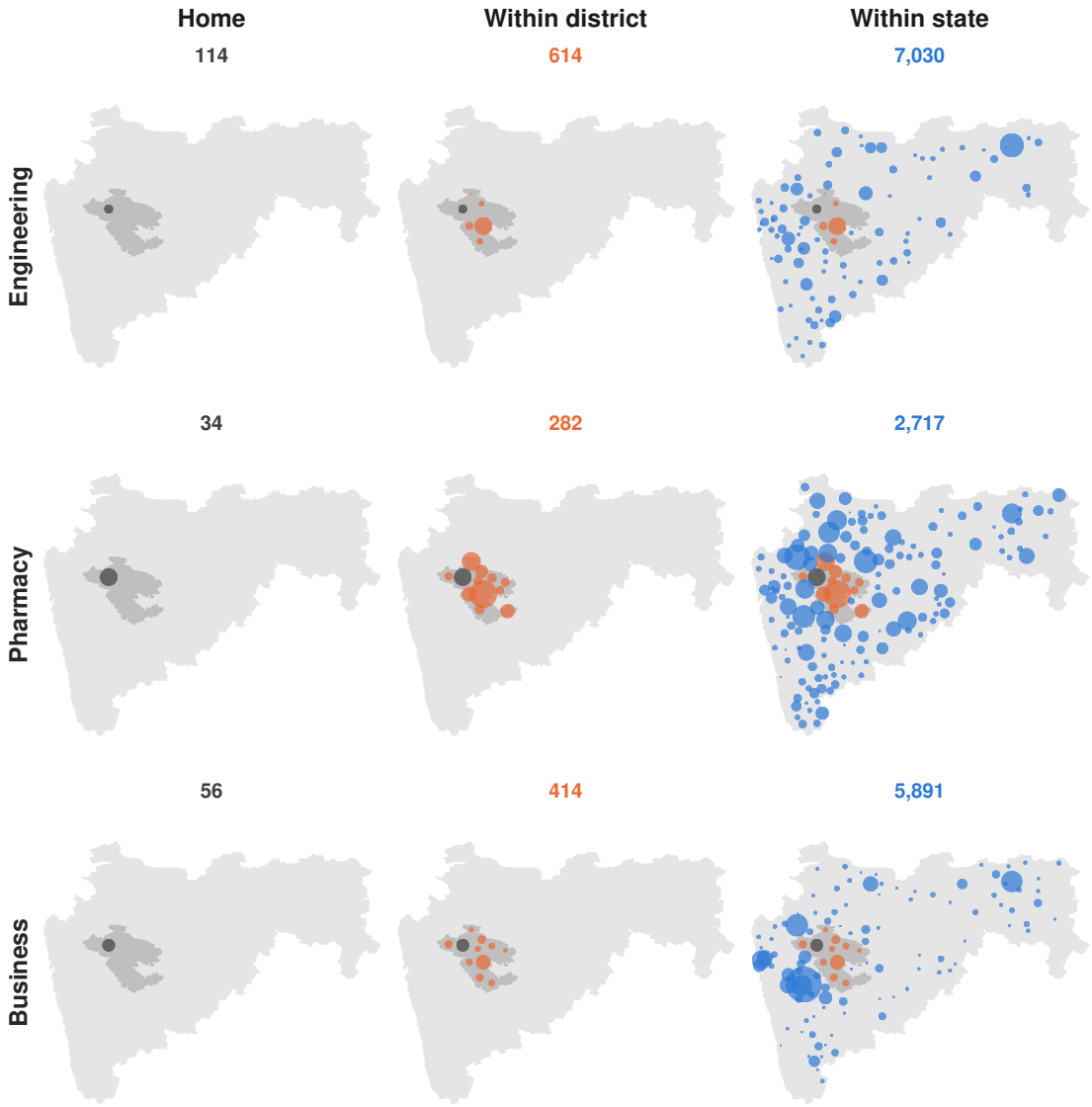


Notes: Panel A counts the institutions in each sub-district that offer programs in each group, and Panel B the students enrolled in 2022–23. The data cover all colleges, including general degree colleges (B.A., B.Com., B.Sc.) that admit students directly. *Science* includes B.Sc., engineering, pharmacy, nursing, medicine and agriculture; *Commerce & Arts* includes B.A., B.Com., business, law and design; *Diploma* covers polytechnic, nursing, pharmacy and teacher-training diplomas. Programs that require a prior degree are excluded. Enrollment in institutions that offer both science and commerce-and-arts programs is split between the two groups. Each map has its own color scale, shown below it. Gray marks sub-districts with no institution in the group; institutions whose address could not be matched to a sub-district hold 7–9 % of enrollment. Source: All India Survey on Higher Education.

Accessible seats across the score distribution. Figure A3 shows the number of accessible seats at every percentile of the entrance exam. The potential gains from moving are large at every point in the distribution. A test-taker at the median has access to about ten times as many engineering seats outside the home district as within it, and a test-taker at the 90th percentile to about twenty times as many.

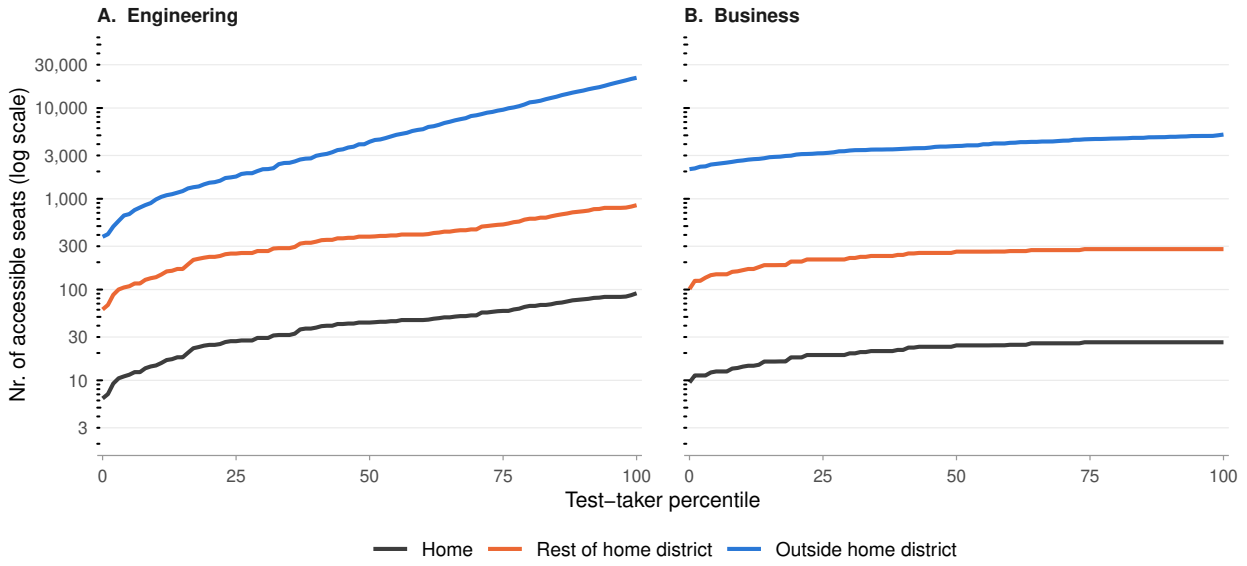
Seats per eligible candidate. The statistics presented so far ignore competition from other candidates. Figure A4 accounts for it by dividing the engineering seats in Ahilyanagar and in Pune by the number of candidates in the state who score at least as high. Both

Figure A2: Accessible seats for the median test-taker, by program and geography



Notes: The numbers above the maps count the 2025–26 seats in the centralized admissions system that a student scoring at the median of the entrance exam could obtain, averaged over all students in the sample. Each student is evaluated at their own home, district, gender and caste category, and the columns are cumulative. A seat counts if its final-round cutoff is at or below the median and the student meets its quota rules; seats still vacant after the final round count at any score, and All-India and minority seats are excluded. The exams are the MHT-CET for engineering and pharmacy and the MAH-BBA/BMS/BBM CET for business. Each map shows one home (dark dot) in one study district (shaded); dots mark places with seats accessible from there, and larger dots mean more seats. Source: State Common Entrance Test Cell, Maharashtra.

Figure A3: Access to engineering and business seats from the study's home sub-districts, by test-taker percentile and geography

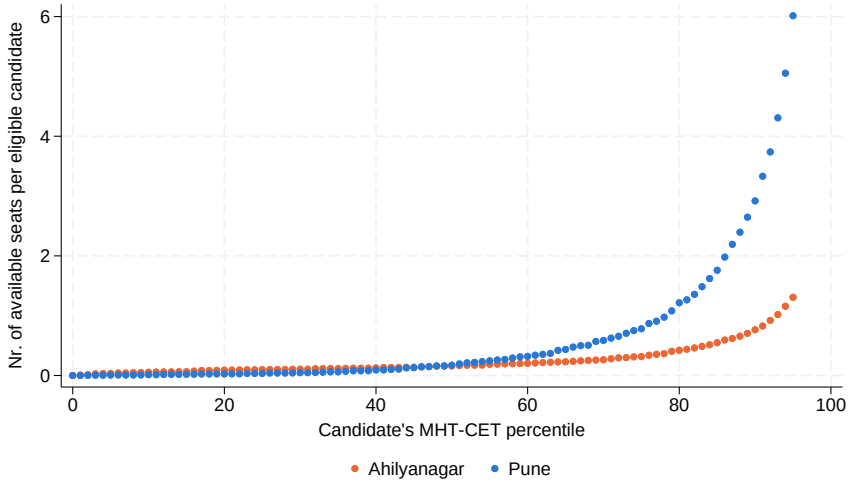


Notes: Each curve counts the 2025–26 seats in the centralized admissions system that a test-taker could obtain at each percentile of the entrance exam, averaged over the 19 sub-districts where the study's schools are located. The test-taker is an open-category man with no special entitlement. The curves separate seats at home, in the rest of the home district and in the rest of Maharashtra. A seat counts if its final-round cutoff is at or below the test-taker's percentile and the test-taker meets its quota rules; seats still vacant after the final round count at any score. Business comprises BBA, BMS and BBM. At the highest score, the averages are 91 engineering seats at home, 850 in the rest of the home district and 21,551 elsewhere (26, 279 and 5,101 in business); 11 of the 19 sub-districts have no accessible engineering seat at home, and 7 none in business. The vertical axis is logarithmic. Source: State Common Entrance Test Cell, Maharashtra.

districts fall under the same home university, so a candidate from Ahilyanagar competes under the same quotas in both. From the 70th percentile up, Pune has two to five times as many seats per eligible candidate as Ahilyanagar, so the gain from moving remains large once competition is taken into account.

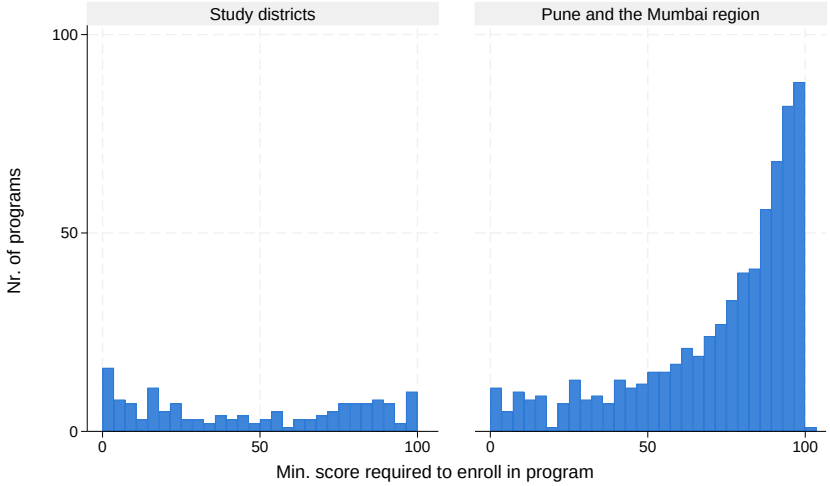
Distribution of admission cutoffs. The figures so far count accessible seats but do not measure their quality. As a revealed-preference measure of program quality, I report each program's admission cutoff for the same type of candidate. Because seats are allocated by entrance-exam score, a program with a higher cutoff is one that more high-scoring candidates ranked above their alternatives. The study districts also skew toward programs with lower admission cutoffs (Figure A5). The right tail of the cutoff distribution (the most selective programs) is concentrated in Pune and the Mumbai region. Moving therefore gives a student access to better programs as well as more of them.

Figure A4: Engineering seats per eligible candidate, MHT-CET 2025–26



Notes: Each point divides two counts at a given MHT-CET percentile. The first is the Maharashtra-quota engineering seats in the district in programs that admitted an open-category man at or below that percentile; the second is the number of open-category men in Maharashtra who scored at or above it. Percentiles above 95 are omitted because few candidates remain. Data are from the 2025–26 final allocation round. Source: State Common Entrance Test Cell, Maharashtra.

Figure A5: Distribution of minimum MHT-CET scores by district



Notes: A program's cutoff is the lowest percentile at which an open-category man was admitted to it in the 2025–26 final allocation round. Each program counts once. The study districts are Ahilyanagar and Sangli; the mentor destinations are Pune and the Mumbai region (Mumbai City, Mumbai Suburban and Thane). Source: State Common Entrance Test Cell, Maharashtra.

A.2 Income gaps

Table A1 reports the income, consumption, and education gaps between Pune and the Mumbai region and the rest of Maharashtra in the India Human Development Survey, the source of the household-survey figures quoted in Section 2. Per-capita incomes in the two regions are 106% higher than in the rest of the state, and 56% higher among households with a graduate, so the regions with the most selective colleges also have higher-paying labor markets.

Table A1: Observational income and education gaps within Maharashtra (IHDS-II)

	Income PC		Consumption PC		Education	
	All HH	HH with graduate	All HH	HH with graduate	Years	Graduate or higher
	(1)	(2)	(3)	(4)	(5)	(6)
Pune and the Mumbai region	1.93***	1.60***	0.96***	0.84***	3.06***	0.10***
	(0.11)	(0.25)	(0.06)	(0.12)	(0.14)	(0.01)
Mean other Maharashtra	1.82	2.88	1.73	2.32	6.80	0.07
N	2,877	592	2,873	592	8,631	8,631

Notes: Each cell reports the coefficient on an indicator for Pune and the Mumbai region from a separate weighted regression on Maharashtra households in the India Human Development Survey (IHDS-II, 2011–12). The Mumbai region comprises Mumbai and Thane. Income and consumption are monthly per-capita values in thousands of 2012 rupees at rural prices, deflated by the official 2011–12 rural and urban poverty lines for Maharashtra recorded in the survey. *Graduate or higher* is at least 15 completed years of education. Columns 1–4 are household regressions with robust standard errors; columns 5–6 are individual regressions for ages 18–64 with age fixed effects and standard errors clustered by household. *Mean other Maharashtra* is the mean outside the two regions. The IHDS is not representative at the district level, so the estimates are descriptive differences. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Appendix: Sample, schools, balance and take-up

This appendix details the restrictions applied to the pool of shortlisted applicants to arrive at the randomization sample (Section 5), and collects the descriptive tables on the schools, the randomization balance, and selection into take-up.

B.1 Sample restrictions

I exclude schools with fewer than five interested applicants for Campus Trail. The sampling and data-collection strategy is designed to recruit students genuinely interested in the program; as an additional data-quality measure, I conduct pre-specified interest and attention checks and keep only students who pass them. Specifically, I exclude students who:

- respond "No" to whether they are interested in applying to Campus Trail both at the beginning and at the end of the survey;
- give at least three of four answers that signal no interest, namely "I want to stop studying after Class 12" when asked about educational plans, "I am not interested

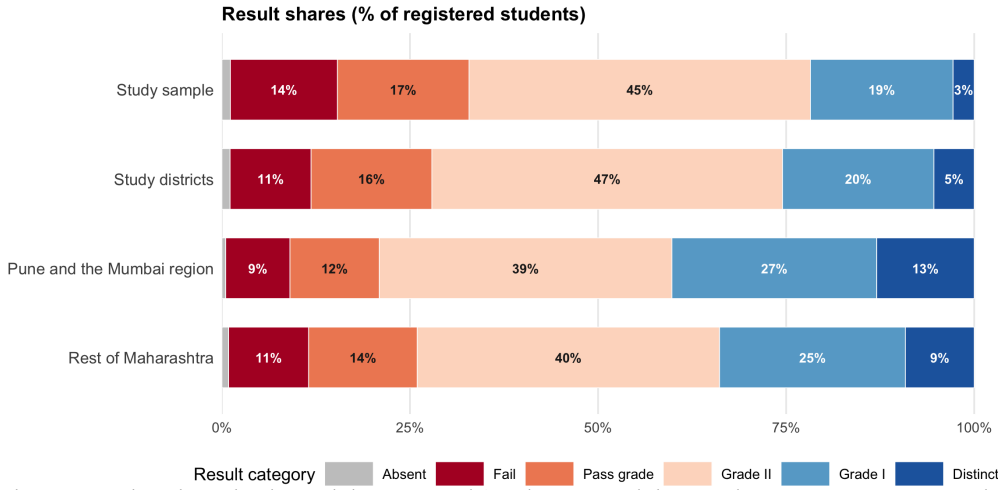
in college recommendations" when asked about potential destinations, "I am not interested in meeting a mentor" and none of the listed topics when asked about topics for mentor guidance; or

- are flagged by field enumerators as inattentive (a score below 3 on the 0–10 attentiveness scale), or explicitly marked as such.

B.2 Schools, balance, and take-up

Appendix Figure B1 compares the 2026 HSC board results of the study-sample schools with the other schools in their districts, Pune and the Mumbai region, and the rest of the state; Table B1 compares the study schools with the other secondary schools in the two study districts; Table B2 reports baseline balance across treatment arms; and Table B3 reports which baseline characteristics predict take-up of mentorship among treated students.

Figure B1: HSC board exam results, 2026: study-sample schools and other schools



Notes: Study districts are the other schools in Ahilyanagar and Sangli. Pune and the Mumbai region are Pune, Mumbai, Mumbai Suburban and Thane. Rest of Maharashtra is every other district.

Table B1: School comparison: study sample vs. other schools in the study districts

	Study sample	Other schools	Study−other
Higher-secondary board: State	0.980	0.954	0.027
	(0.140)	(0.211)	(0.021)
Rural school	0.804	0.683	0.121**
	(0.401)	(0.466)	(0.059)
Marathi medium	0.902	0.765	0.137***
	(0.300)	(0.424)	(0.045)
Has internet	0.941	0.943	-0.002
	(0.238)	(0.232)	(0.034)
Has electricity	1.000	0.996	0.004*
	(0.000)	(0.067)	(0.003)
Private school	0.196	0.364	-0.168***
	(0.401)	(0.481)	(0.059)
No. of 12th-grade students	127.2	132.2	-4.9
	(94.0)	(178.4)	(14.8)
Share girls in 12th grade	0.486	0.411	0.074***
	(0.147)	(0.229)	(0.022)
Teachers per student	0.029	0.056	-0.028***
	(0.009)	(0.141)	(0.006)
Classrooms per student	0.025	0.091	-0.065***
	(0.010)	(0.461)	(0.018)
Projectors per student	0.005	0.013	-0.008***
	(0.005)	(0.062)	(0.003)
Joint F-test p			<0.001
Schools	51	668	719

Notes: Study-sample schools versus all other secondary schools in the two study districts (Ahilyanagar and Sangli), from UDISE administrative school records. One study-sample school is unmatched in these records, so the comparison covers 51 of the 52 study schools. Cells report group means with standard deviations in parentheses. The last column reports the study-sample minus other-schools difference with robust standard errors in parentheses, and the joint F-test asks whether the characteristics jointly predict being a study school. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B2: Summary statistics and randomization balance

	Control	Treated	T−C
Female	0.617	0.599	-0.018
	(0.486)	(0.490)	(0.024)
Age	17.221	17.259	0.038
	(0.617)	(0.637)	(0.031)
General category	0.367	0.355	-0.012
	(0.482)	(0.479)	(0.023)
Past board marks	71.005	70.795	-0.210
	(11.334)	(10.926)	(0.543)
Household size	5.424	5.383	-0.041
	(2.268)	(2.206)	(0.109)
Asset index	0.000	0.028	0.028
	(1.000)	(1.151)	(0.053)
Has electricity	0.971	0.956	-0.014
	(0.168)	(0.204)	(0.009)
Tap water at home	0.508	0.535	0.027
	(0.500)	(0.499)	(0.024)
Pukka house	0.647	0.684	0.037
	(0.478)	(0.465)	(0.023)
Mother completed class 10	0.537	0.519	-0.018
	(0.499)	(0.500)	(0.024)
Father completed class 10	0.687	0.697	0.010
	(0.464)	(0.460)	(0.023)
Travels once a year or more	0.748	0.723	-0.025
	(0.435)	(0.448)	(0.022)
Internet access	0.826	0.817	-0.009
	(0.379)	(0.387)	(0.019)
Self-efficacy	16.218	16.424	0.205
	(5.032)	(4.727)	(0.238)
Joint orthogonality F-test p			0.400
Observations	856	826	1,682

Notes: Baseline means by treatment arm for the analysis sample, with standard deviations in parentheses. The last column reports the treated-minus-control difference with robust standard errors in parentheses, together with the p-value of the joint orthogonality F-test.

Table B3: Selection into take-up

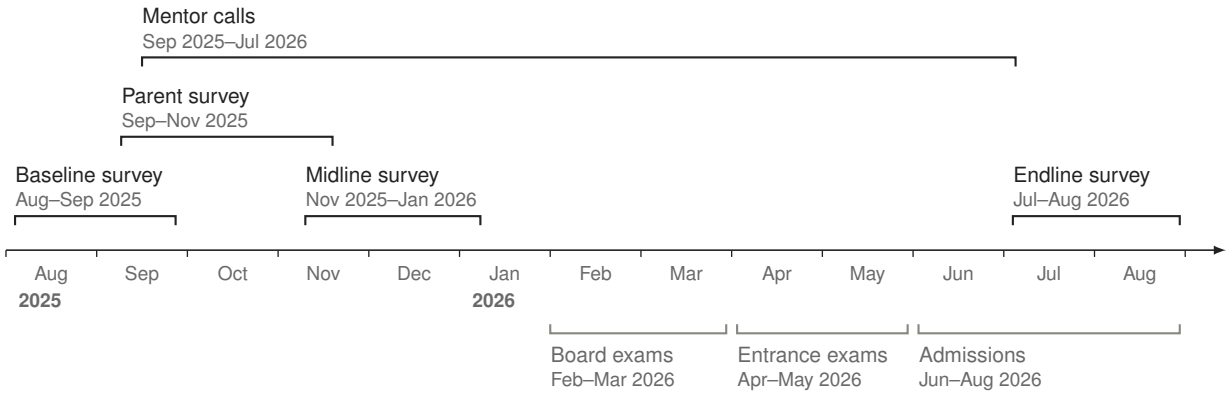
	No take-up	Took up	Took up−no
Female	0.619	0.576	-0.043
	(0.486)	(0.495)	(0.034)
Science stream	0.650	0.734	0.084***
	(0.477)	(0.442)	(0.032)
Age	17.334	17.171	-0.163***
	(0.631)	(0.633)	(0.044)
General category	0.341	0.371	0.030
	(0.475)	(0.484)	(0.033)
Past board marks	69.753	72.018	2.264***
	(10.914)	(10.827)	(0.759)
Expected board marks	72.227	73.303	1.076
	(9.628)	(9.846)	(0.680)
Asset index	0.076	-0.028	-0.104
	(1.152)	(1.148)	(0.080)
Father completed class 10	0.697	0.697	0.000
	(0.460)	(0.460)	(0.032)
Internet access	0.805	0.832	0.027
	(0.397)	(0.375)	(0.027)
Household size	5.516	5.226	-0.289*
	(2.315)	(2.063)	(0.152)
Willingness to take risk	5.430	5.797	0.367*
	(3.141)	(3.083)	(0.217)
Number of close friends	4.545	4.424	-0.121
	(3.501)	(3.343)	(0.239)
Self-efficacy	15.982	16.942	0.960***
	(4.679)	(4.737)	(0.329)
Has decided on a plan	1.020	1.082	0.061
	(0.910)	(0.890)	(0.063)
Any network away from home	0.558	0.576	0.018
	(0.497)	(0.495)	(0.035)
P(apply) away from home	0.413	0.487	0.075***
	(0.284)	(0.287)	(0.020)
Rating of away destinations	5.656	6.042	0.386*
	(2.952)	(2.865)	(0.203)
Social desirability (M–C)	6.827	6.858	0.031
	(2.312)	(2.350)	(0.163)
Effort index	-0.019	0.033	0.052
	(1.003)	(0.985)	(0.069)
Knowledge index	-0.047	0.021	0.068
	(0.903)	(1.055)	(0.069)
Joint F-test p			<0.001
Treated students	446	380	826

Notes: Baseline characteristics of treated students who took up mentorship (at least fifteen minutes of recorded conversation) versus those who did not. Columns report group means with standard deviations below; the last column reports the took-up minus did-not-take-up difference from an OLS regression, with the robust standard error below, and the joint-test row reports the p-value of a joint F-test that the characteristics jointly predict take-up. The covariate set is the one used to predict take-up in the causal-forest analysis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C Appendix: Study timeline

Figure C1 shows when each survey wave and the mentor calls took place, relative to the exam and admissions calendar. The baseline and midline surveys took place in schools before the board and entrance exams, and the endline survey took place by phone as admissions concluded.

Figure C1: Study timeline



Notes: Brackets run from the first to the last interview of each survey and from the first to the last recorded mentor call. For each treated student, mentor calls started only once the parent survey was completed or all attempts to reach a parent were exhausted, so the two periods overlap only across students. Exam and admission periods are approximate.

D Appendix: Mentor applicant pool and migrant-peer evidence

This appendix describes how mentors are recruited and the pool of students who applied to be Campus Trail mentors. These data are not meant to be representative of educational migrants, nor of the mentors who ultimately worked in the program. Applicants selected into the form because they were interested in mentoring, and the experiences survey, a follow-up questionnaire about their own move (their studies, costs and financial aid, contacts in the city, and life after moving), was completed by a smaller subset of this pool. I therefore use the evidence here descriptively, as anecdotal information on what kinds of destination-specific experience the program could transmit to mentees.

Recruitment. Mentors are recruited with a flyer, shared in college WhatsApp groups and other informal networks (Figure D1), and through the program website. The flyer invites university students and recent graduates who moved to Pune or Mumbai from another city in India to study to apply as mentors, called Trailblazers. It describes the role as having conversations with prospective students about university life in a new city,

including housing, culture, challenges, and opportunities. It advertises payment of up to INR 800 per session, leadership experience, and a stronger CV, and it states that Campus Trail is part of a research study. The website lists three eligibility conditions. Applicants must have moved to Pune or Mumbai for university from another city or town in India, must have completed at least one academic year, and must be comfortable speaking about their experience and answering students' questions. Applicants fill in an online application form with their CV and a recent marksheet, are interviewed by the program team, complete an online training session, and prepare for their sessions. They may also complete the experiences survey. Mentors receive a recommendation letter in addition to their pay.

Figure D1: Mentor recruitment flyer



Applicant profile. Table D1 summarizes the applicant pool characteristics. Applicants are older and academically selected relative to the high-school students in the experi-

ment. Most study in Pune or Mumbai and most report having moved specifically for studies, usually alone. Among survey respondents, 40 % receive some financial assistance for their studies, mostly scholarships. They report substantially higher living costs compared to the student sample. Given their selection into migration and into applying as mentors, I do not interpret these as a representative estimate of what the marginal student in the experiment would spend if they moved.

Table D1: Summary statistics for the mentor applicant pool

Variable	N	Mean	SD	Median
Application form applicant pool				
Female	83	0.59	0.49	1.00
Current undergraduate	83	0.53	0.50	1.00
Studying in Pune/Mumbai region	83	0.88	0.33	1.00
Moved specifically for studies	83	0.78	0.41	1.00
Moved alone	83	0.78	0.41	1.00
Knew someone in destination	83	0.53	0.50	1.00
12th-board percentage	78	79.26	12.68	80.50
Experiences survey respondent pool				
Monthly expenses (Rs.)	33	14,948	10,257	13,000
Annual tuition (Rs.)	33	75,047	79,090	40,000
Accommodation plus food (Rs./month)	33	16,369	18,435	10,000
No financial assistance	35	0.60	0.50	1.00

Notes: The application-form panel uses applicant records with a non-missing institution name. The experiences-survey panel uses respondents with a non-missing institution name. Indicator means are shares within this selected pool. Monetary variables are cleaned by dropping negative values and values of Rs. 1,000,000 or more, which are treated as invalid/sentinel entries. The table is descriptive only; the applicant pool is not representative of educational migrants or of the mentors who ultimately worked in the program.

Applicants' own contacts at the destination. The experiences survey asks migrant applicants whether they had contacts in the destination before moving, what help those contacts provided, and what help would have been useful if they had no contacts. Across both groups, the recurring themes are practical and destination-specific: what living in the city costs, where a student can live safely and affordably, which colleges or programs have a good reputation, what campus life is like, and how to manage the first weeks after arrival.

What applicants propose to provide to mentees. The application form also asks applicants which topics they would discuss in a one-hour conversation with a 12th-grade

student considering study in their city. In addition, each applicant receives a randomized vignette describing one concern a hypothetical mentee has: costs and housing, academic returns and jobs, missing home and making friends, family and community pressure, daily adjustment, or safety and navigation, and is asked how they would advise the student. The baseline survey asked students a similar question about which topics they would raise with a current student in a destination city. Their most common choices were cost of living, career prospects, safety and academics; adjusting to university life was the least common.

Analyzing the responses offers rich insights into how the mentor pool views the transition into college life, and what type of information, support or encouragement could flow to mentees. When the vignette is about costs and housing, applicants often translate the concern into a concrete budget: rent ranges, food costs, transport passes, hostels, paying-guest rooms, tiffin services, scholarships, educational loans, and sometimes part-time work. When the vignette is about academic returns, they tend to compare the destination option with staying closer to home, discussing course structure, faculty, placements, alumni, internships, and the skills students would need beyond the degree itself. When the concern is family or community pressure, applicants often switch to a parent-facing argument: moving is framed as an investment in career prospects, exposure, independence, and long-run family benefit. For homesickness and daily adjustment, applicants discuss loneliness, describe routines for managing cooking, budgeting, and chores, and point to roommates, seniors, clubs, student groups, professors, and counseling services as sources of support. For safety and navigation, they usually combine reassurance with concrete precautions: safe areas, transport options, emergency contacts, campus security, and people the student could call after arrival.

E Appendix: Destination options by survey wave

Several survey modules ask students to report beliefs, preferences, information, or hypothetical choices for named destinations. Besides home, the destinations are Pune, Mumbai and Nagpur, the three main higher-education destinations in Maharashtra, and from midline on the student's district capital. Each student evaluated home and a random subset of these destinations (Table E1), and the outcomes compare each away destination with home.

After baseline, I refined the definition of home from the district to the sub-district, a change documented in the midline pre-analysis plan. During baseline fieldwork, the field team observed that many students answered the district questions with the district

capital in mind, so the answers did not measure home for students living elsewhere in the district. From midline on, the modules therefore ask about the student's own sub-district and treat the district capital as a separate destination. Baseline outcomes compare destinations outside the district with the district, and midline and endline outcomes compare each away destination, including the district capital, with the sub-district.

Estimates that pool baseline and midline adjust for survey wave. In the pooled belief figures (Figures 3 and 4), baseline answers about the district are shown as home for students who live in the district capital and at the district capital for students who live elsewhere; Appendix F describes how the pooled choice experiment treats them. The choice-experiment estimates are also reported for each wave separately (Table F4), and the endline outcomes for moves out of the home district (Table L3).

Table E1: Destination options used in destination-specific survey modules

Survey wave	Destination pool	Destinations each student evaluated
Baseline	Own district; Pune; Mumbai; Nagpur.	Each student evaluated three destinations, their own district and two of Pune, Mumbai and Nagpur. The first was randomized between Pune and Mumbai, the second between Nagpur and the other of Pune and Mumbai.
Midline	Home (own sub-district); district capital (Ahilyanagar city for students from Ahilyanagar district, Sangli-Miraj for students from Sangli district); Pune; Mumbai.	Each student evaluated three destinations, home and two others. One was always the student's mentor city. Treated students had their actual mentor city, and control students had a pseudo-mentor city assigned by the same matching procedure. For students living outside the district capital, the other was randomized between the district capital and the other of Pune and Mumbai. Students living in the district capital evaluated Pune and Mumbai.
Endline	Home; district capital; Pune.	Each student evaluated two destinations, home and one comparison destination assigned before endline. Students living in the district capital were assigned Pune. Other students were randomized to Pune (one third) or the district capital (two thirds), within cells of treatment, gender and stream. Beliefs were first asked for the student's realized or planned destination, which could lie outside this set, and then for home and the comparison destination where still missing.

F Appendix: Discrete choice experiment

This appendix documents the construction and presentation of the DCE attributes described in Section 4.1.

Overview. In both survey waves, each choice task presented two hypothetical college profiles and a "No university" opt-out. The randomized profiles varied in destination and in one return or quality attribute, while the remaining profile details were fixed at the student level after being constructed from the student's own survey responses. The

designs were generated in Stata using the `dcreate` module for efficient discrete-choice designs (Hole, 2015).

Destination sets. The DCE used the student-specific destination set constructed for the relevant survey wave. Appendix E describes the baseline and midline destination pools and selection rules.

Profile construction. Several profile values were anchored to earlier survey responses, while others were fixed at the student level to keep the tradeoff focused on destination and the arm-specific return or quality measure. Table F1 summarizes both the anchor source and the displayed construction for each profile item.

Table F1: DCE profile attributes and anchors

Attribute	Fixed or randomized	Construction
Field of study	Fixed at student level	"What would you like to study at university?"; if multiple fields were selected, the survey asked for the most preferred field. All profiles shown to the student were labeled as colleges in that field.
Destination	Randomized	Randomized within the student-specific destination set constructed for the wave; see Appendix E.
Educational quality (one attribute per arm)		
Expected earnings (baseline and midline)	Randomized	Anchor: "How much money do you think you will earn per year at age 25?" The survey elicited a bracket and an exact amount; the DCE used the bracket anchor. Levels: 0.75, 1.00, 1.25, or 1.50 times the bracket anchor, converted to rupees.
Employment probability (baseline only)	Randomized	Anchor: 0–10 response to "How likely is it that you will be working at a full-time job when you are 25 years old?", bounded at 2–8 and multiplied by 10. Levels: anchor -10, anchor, anchor $+10$, or anchor $+20$ percentage points.
Admission cutoff (baseline only)	Randomized	Anchor: expected entrance-exam score when relevant; otherwise expected 12th-board score, bounded between 20 and 80. Program-level MHT-CET data were used to convert expected scores into cutoff scores corresponding roughly to program-rank movements. Levels: cutoff scores out of 100, corresponding roughly to 50 ranks better (a lower rank is a more selective program), the reference cutoff, or 50 ranks worse. A coding error replaced the intended fourth level, 100 ranks better, with the reference cutoff.
Yearly college fees	Fixed at student level	Family budget for yearly college fees. Shown: 70 percent of the lower endpoint of the selected budget bracket; Rs. 3,000 if the bracket minimum was zero.
Monthly room and food costs	Fixed at student level	Family budget for monthly living expenses while studying in college, including room, food, and other expenses. Shown: 90 percent of the lower endpoint of the selected budget bracket; Rs. 2,000 if the bracket minimum was zero.

Baseline design. The baseline DCE has three randomized arms: expected earnings, employment probability, and admission cutoff. In each arm, a profile combines one of three destinations with one of four levels of the arm's quality attribute (Table F1).

For each arm, I generated 70 D-efficient two-profile choice sets from these twelve profiles using uninformative priors and blocked them into ten seven-question blocks. Students were assigned to one of the resulting 30 blocks. Motivated by Clayton et al. (2026), I then added an eighth task to each block by repeating the second task with the left and right profiles mirrored. This mirrored pair provides a check of response consistency (Appendix F.2). Within each block, both question order and profile order were randomized. The opt-out was displayed in the baseline survey as a third response option alongside the two randomized college profiles.

Midline design. The midline DCE focuses on the earnings arm and uses priors estimated from the baseline earnings-arm choices. Specifically, I estimate conditional-logit priors using baseline respondents who gave consistent answers on the mirrored baseline pair, and then use those priors to generate a Bayesian D-efficient design. The midline design includes three destinations and the same four earnings levels as the baseline earnings arm. The design contains 80 choice sets blocked into ten eight-question blocks, with question order and profile order randomized within block.

Pooled estimation sample. The estimates in Table 1 pool the baseline earnings-arm choice sets with the midline choice sets of control students, so that the district capital, offered only at midline, can be placed on the same ladder as the out-of-district cities. The estimating equation is

$$U_{ijt} = \sum_{g} \beta_{0g} \text{NoUniversity}_j \, \mathbf{1}\{g_i = g\} + \sum_{d} \delta_d \, D_{dj} \, (1 - B_{it}) + \sum_{d} \gamma_d \, D_{dj} \, B_{it} + \boldsymbol{\lambda}' \mathbf{S}_j + \varepsilon_{ijt}, \tag{5}$$

where g indexes four groups of students (baseline or midline, crossed with living in a district capital or not), D_{dj} indicates that alternative j is in destination d (the district capital, Pune, Mumbai, or Nagpur), $\mathbf{S}_j$ collects the salary levels relative to the expected level, and B_{it} equals one for the baseline choices of students who live outside a district capital. For baseline students who live in a district capital the own-district option is their home, and their Pune, Mumbai, and Nagpur options identify those destination coefficients; baseline students who live elsewhere in the district saw their district as a whole as the home option, so their destination comparisons are not relative to home. They enter with their own destination coefficients γ_d, which are not reported, and contribute to the salary coefficients $\boldsymbol{\lambda}$. At midline, students living outside a capital identify the district-capital coefficient, and all control students contribute to the Pune and Mumbai coefficients. The salary levels are common across waves, each wave and residence group has its own opt-

out constant, and standard errors are clustered by student. The main estimates use every completed choice task (981 students, 555 at baseline and 614 at midline, 9352 choice sets); the attentive-sample row of Table F3 repeats the estimates of migration costs of leaving home after dropping, in both waves, students who chose the same option position in every task, the pre-specified attention check. The belief-controlled estimates add the eight destination-specific beliefs of Section 4.1, measured in the wave of the choice task (the baseline city module for baseline tasks, the midline city module for midline tasks), with moving difficulty set to its definitional home value in both waves; each wave's beliefs enter with their own coefficients. The no-university alternative carries the reference level of the randomized attributes and, in the belief-controlled models, the student's beliefs about home, under the assumption that not enrolling means staying at home; Appendix Table F4 reports the alternatives.

Presentation. Before the baseline choice tasks, students saw the note: "Please click on the "Play Button" to watch the video for instructions for the next section!" The accompanying instruction image began: "Which college would you choose? You're about to see information about pairs of imaginary colleges. In each pair, imagine you've scored high enough to get admission into both." It then explained that, for each college, the survey would show living costs, college fees, location, and the arm-specific quality measure: "placement packages" in the earnings arm, "share of students who got jobs through campus placement" in the employment arm, and "cutoffs" in the cutoff arm. The screen listed the four profile details and then stated: "The only differences are the 4 details we'll show you. Everything else about the colleges is the same!" It ended: "In each round, you'll be given only three options: College 1, College 2, or not going to college at all. Which one would you choose? There's no right or wrong answer – just what feels perfect for your future!" Figure F1 shows the instruction poster and an example choice task.

(a) Instruction poster



(b) Choice task screen

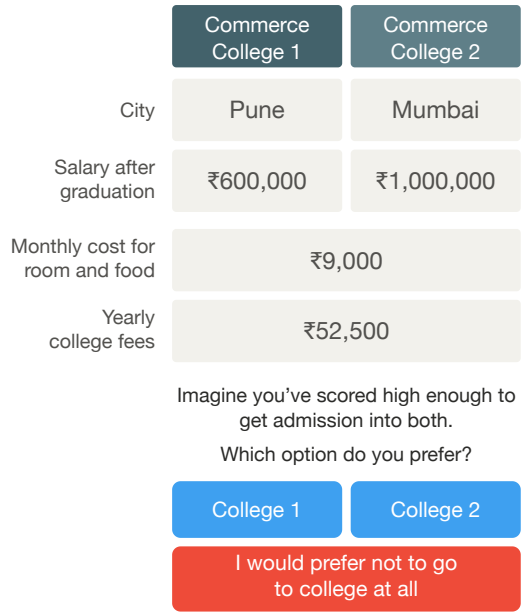


Figure F1: DCE instruction and choice-task screens shown in the baseline survey

F.1 Additional results

Table F2 reports migration costs by destination, in percent of expected salary and in dollars. Table F3 reports migration costs of leaving home in each baseline arm estimated on its own and, for the earnings arm, pooled with the midline control group, for all students, by gender, and on the attentive sample. Table F4 shows how much of the disutility of leaving home beliefs account for, across samples, and Table F5 splits it by belief.

Table F2: Migration costs by destination

	District capital	Pune	Mumbai	Nagpur
Panel A. Utility coefficient on the destination (home = 0)				
All	-0.176**	-0.423***	-0.893***	-1.379***
	(0.080)	(0.060)	(0.072)	(0.208)
Men	-0.219*	-0.482***	-0.709***	-1.604***
	(0.118)	(0.094)	(0.117)	(0.371)
Women	-0.163	-0.401***	-0.997***	-1.287***
	(0.109)	(0.078)	(0.092)	(0.249)
Panel B. Migration costs, % of expected salary				
All	39**	93***	196***	303***
	(19)	(22)	(39)	(69)
Men	28*	63***	92***	209***
	(16)	(16)	(22)	(58)
Women	60	149**	370**	478**
	(48)	(67)	(149)	(203)
Panel C. Migration costs, thousand dollars of yearly salary				
All	1.79**	4.19***	8.81***	13.52***
	(0.83)	(0.79)	(1.34)	(2.68)
Men	1.92	4.60***	6.83***	15.36***
	(1.20)	(1.19)	(1.65)	(4.45)
Women	1.73	4.05***	10.01***	12.92***
	(1.15)	(1.10)	(2.12)	(3.53)

Notes: Conditional logit estimates from the choice experiment, with one indicator per destination and home as the reference. Standard errors clustered by student in parentheses, computed by the delta method for migration costs. The sample pools the baseline earnings arm with the midline control group (981 students: 365 men and 616 women). Panel A reports the destination coefficients from the model with salary in percent of the student's expected salary. Migration costs are the destination coefficient divided by the salary coefficient of the same model, with salary in percent of expected salary (Panel B) or in thousand dollars per year (Panel C, at 93.75 INR per USD). Men and women are estimated separately. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F3: Migration costs by numeraire arm and sample

	Earnings arm				Employment arm	Cutoff arm
	% of expected salary		Thousand dollars of yearly salary		Percentage points	Program ranks
	Baseline	Pooled	Baseline	Pooled	Baseline	Baseline
	(1)	(2)	(3)	(4)	(5)	(6)
Utility coefficients						
Away from home	-0.774***	-0.587***	-0.795***	-0.595***	-0.847***	-0.740***
	(0.061)	(0.054)	(0.062)	(0.055)	(0.064)	(0.065)
Numeraire (one unit)	0.005***	0.005***	0.129***	0.103***	0.027***	0.003***
	(0.001)	(0.001)	(0.020)	(0.014)	(0.003)	(0.001)
Migration costs: away coefficient / numeraire coefficient						
All	142***	129***	6.17***	5.84***	31.4***	253***
	(33)	(26)	(1.04)	(0.92)	(4.2)	(64)
Men	82***	72***	6.05***	5.26***	27.3***	271*
	(18)	(16)	(1.38)	(1.21)	(5.9)	(150)
Women	267*	227**	6.36***	6.33***	34.4***	244***
	(141)	(93)	(1.63)	(1.42)	(6.0)	(67)
Attentive sample	126***	123***	6.22***	6.15***	31.1***	231***
	(27)	(24)	(1.04)	(0.96)	(4.1)	(53)
Students	555	981	555	981	575	552

Notes: Conditional logit estimates from the choice experiment, with one indicator for any destination away from home. Standard errors clustered by student in parentheses, computed by the delta method for migration costs. *Baseline* columns use the baseline survey only, with the student's district as home. *Pooled* columns add the midline control group, which saw only the earnings arm. Each arm varies one attribute around the student's expected level: earnings by -25, $+25$ and $+50$ percent, the probability of employment by -10, $+10$ and $+20$ percentage points, and admission cutoffs by ± 50 program ranks. A display error in the survey software showed the reference cutoff in place of the fourth cutoff level; these alternatives are coded at the reference level. Columns 3 and 4 measure salary in thousand dollars per year instead of percent, at 93.75 INR per USD. The numeraire coefficient is the coefficient on a one-unit increase, averaged over the levels shown. Migration costs are the away coefficient divided by the numeraire coefficient, in the units of the column. Men and women are estimated separately. The attentive sample drops students who chose the same option position in every choice, the pre-specified attention check. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F4: How much of the disutility of leaving home beliefs account for

	All	Men	Women
Panel A. Migration costs of leaving home, % of expected salary			
Without belief controls	129***	72***	227**
	(26)	(16)	(93)
With belief controls	42**	63***	4
	(18)	(18)	(38)
Reduction (%)	67.3	11.7	98.3
Panel B. Share of the away coefficient accounted for by beliefs (%)			
All arms pooled with the midline control group	61.3	41.6	74.1
Changing the sample			
Baseline only, earnings arm	35.9	23.2	45.3
Baseline only, employment arm	60.0	51.8	66.6
Baseline only, cutoff arm	75.1	66.7	80.9
Midline control group only	76.2	18.1	104.2
Leaving out moving difficulty			
All arms pooled with the midline control group	34.8	28.0	39.4
Baseline only, earnings arm	24.4	20.5	27.8
Baseline only, employment arm	37.7	35.7	40.3
Baseline only, cutoff arm	42.9	41.1	43.7
Midline control group only	33.4	6.3	46.1
Assigning not studying the colleges' average beliefs			
All arms pooled with the midline control group	46.2	39.9	48.9

Notes: Each row compares a model of choice with an indicator for any destination away from home, without and with the eight beliefs of Section 4.1 as controls, on choices with complete belief data. Panel A uses the conditional logit of the main estimates, pooling the baseline earnings arm with the midline control group (981 students); standard errors clustered by student in parentheses, computed by the delta method. Panel B uses a linear probability version with choice-set fixed effects and reports the percentage by which the away coefficient falls once beliefs are added (Gelbach, 2016). Its first row pools all three baseline arms with the midline control group (1682 students), with the student's district as home at baseline. The indented rows change one thing at a time: the sample, the beliefs included, and the beliefs assigned to the option of not studying, which are beliefs about home in every other row. Shares above 100 mean that the away coefficient changes sign. Men and women are estimated separately. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F5: Share of the disutility of leaving home accounted for by each belief

	All	Men	Women
Emotional well-being	13.3	16.2	12.5
95% CI	[10.0, 16.7]	[9.6, 23.1]	[8.9, 16.3]
Happiness	12.6	16.0	11.3
Mental wellbeing	0.7	0.2	1.2
Social ties	14.1	11.7	15.1
95% CI	[10.7, 18.0]	[6.4, 17.9]	[10.1, 20.5]
Reputation back home	0.4	0.2	0.5
Days seeing family	0.9	1.0	0.7
Close friends at destination	2.3	2.0	2.5
Parental approval	10.5	8.6	11.4
Practical and lived experience	34.0	13.7	46.5
95% CI	[24.1, 44.6]	[-2.0, 30.3]	[33.5, 59.7]
Safety	4.2	-1.6	7.4
Moving difficulty	29.8	15.3	39.1
Total	61.3	41.6	74.1
95% CI	[51.6, 72.4]	[27.8, 58.6]	[61.1, 88.0]
Away coefficient, without beliefs	-0.279***	-0.246***	-0.300***
	(0.015)	(0.023)	(0.020)
Away coefficient, with beliefs	-0.108***	-0.144***	-0.078***
	(0.021)	(0.031)	(0.028)
Students	1,682	659	1,023

Notes: Linear probability model of choice on an indicator for any destination away from home, with choice-set fixed effects, pooling all three baseline arms with the midline control group, on choices with complete belief data. Standard errors clustered by student in parentheses. Each belief's share is its coefficient in the model with beliefs times the average gap in that belief between the away option and home within a choice, summed over the two survey waves and expressed as a percentage of the away coefficient without beliefs (Gelbach, 2016). The shares add up to the total. Family subtotals and the total carry 95% confidence intervals from a bootstrap that resamples students within gender (1000 replications). Men and women are estimated separately. Beliefs are positively correlated and the joint model divides their shared variation among them, so a single belief's share understates what it accounts for on its own.

F.2 Inattention and measurement error

Random inattention in conjoint experiments attenuates estimated effects toward zero (Clayton et al., 2026). Following their design recommendation, the second and eighth choice tasks mirror each other: the same profiles reappear with positions swapped. A respondent answering attentively should choose the same profile both times.

Between 75 and 79 % of students do (Table F6). Under the benchmark model in which a respondent answers randomly with probability 2ϵ (so two independent measurements of the same choice disagree with probability $2\epsilon(1-\epsilon)$), the implied swapping error is $\epsilon = 0.12$–0.14, and observed effects understate true effects by a factor of $1-2\epsilon \approx 0.71$–0.76. The table reports three checks. First, the away effect on the choice probability, rescaled by $1/(1-2\epsilon)$, rises by about a third in magnitude. The correction scales all attribute effects equally, so estimates of migration costs are unaffected by this form of error. Second, re-estimating on the consistent respondents only raises the away coefficient by 12 to 18 percent, as attenuation predicts, and leaves the estimates of migration costs close to those on the attentive sample in Table F3. Third, dropping the repeated task changes the away coefficient by less than one percent. Migration costs are therefore not an artifact of inattentive responding.

Table F6: Inattention in the DCE: mirrored-task consistency and corrections

	Earnings	Employment	Cutoffs
Consistent on mirrored task (IRR)	0.777	0.788	0.754
Implied swapping error ϵ	0.128	0.120	0.144
Implied attenuation $1-2\epsilon$	0.744	0.759	0.712
Away effect on choice (AMCE)	-0.136	-0.176	-0.022
	(0.026)	(0.026)	(0.024)
AMCE corrected by $1/(1-2\epsilon)$	-0.183	-0.232	-0.031
	(0.034)	(0.034)	(0.034)
Away coefficient, consistent only	-0.918	-1.110	-0.901
	(0.084)	(0.088)	(0.092)
MRS to stay home, consistent only	136	32	206
	(34)	(4)	(46)
Away coefficient, excl. repeated task	-0.822	-0.945	-0.806
	(0.069)	(0.072)	(0.074)
Nr. students (consistent / all)	380 / 489	402 / 510	358 / 475

Notes: IRR = share of students choosing the same profile in the mirrored second and eighth tasks. The swapping error solves $2\epsilon(1-\epsilon) = 1 - \text{IRR}$, treating the three-option choice as binary between the two college profiles. The average marginal component effect (AMCE) is the coefficient on the away indicator from a linear probability model of choice on the attributes with choice-set fixed effects; the corrected row rescales it by $1/(1-2\epsilon)$ per Clayton et al. (2026). Consistent-only rows re-estimate the conditional logit of each baseline arm on students who passed the mirror check; the final coefficient row drops the repeated (eighth) task. Standard errors clustered by student.

G Appendix: College rating exercise

This appendix documents the college rating exercise, which measures how students value college location when profile attributes are varied directly. Students rated hypothetical college profiles in the baseline and midline surveys, three profiles at baseline and five at midline. The willingness-to-pay benchmark reported in Section 4.2 uses the baseline wave.

Overview. Students rated the profiles one at a time. Each profile was described as a college in the student's most-preferred field. Each screen asked how likely the student would be to accept admission if the displayed college were the only option. Responses were recorded on a 0–10 scale.

Profile construction. Each profile was generated inside the survey form using fresh random draws. The destination was drawn from the same destination cities used throughout the survey (Appendix E). Fees, living costs, hostel information, admission cutoff,

placement share, average earnings, and higher-education share were randomized at the profile level (Table G1). Earnings and higher-education shares were not drawn independently of the rest of the profile: profiles with higher placement shares tended to have higher earnings and lower higher-education shares, with additional random variation. These correlations made the profiles look more like realistic college information while preserving experimental variation in each attribute.

Table G1: College rating profile attributes

Profile item	Fixed or randomized	Construction
Field of study	Fixed	Profiles were labeled as colleges in the student's most-preferred field.
Destination	Randomized	One of the destination cities used throughout the survey.
Yearly college fees	Randomized	Rs. 1,000 plus a uniform draw between Rs. 0 and Rs. 100,000, rounded to the nearest rupee.
Monthly food and living costs	Randomized	Rs. 5,000 plus a uniform draw between Rs. 0 and Rs. 50,000, rounded to the nearest rupee.
Hostel information	Randomized	Four equally likely categories: clean and safe college hostel; college hostel available but quality uncertain; only private hostel or PG options available; and no information.
Admission cutoff	Randomized	A uniform draw between 40 and 100, displayed as the cutoff out of 100 for the general category.
Job placement share	Randomized	A uniform draw between 0 and 100, displayed as the share of students placed in jobs after graduation.
Average yearly earnings	Randomized	Generated as a positive function of the placement share plus random noise, with a lower bound of Rs. 100,000 per year.
Higher-education share	Randomized	Generated as a negative function of the placement share plus random noise, bounded between 0 and 100.

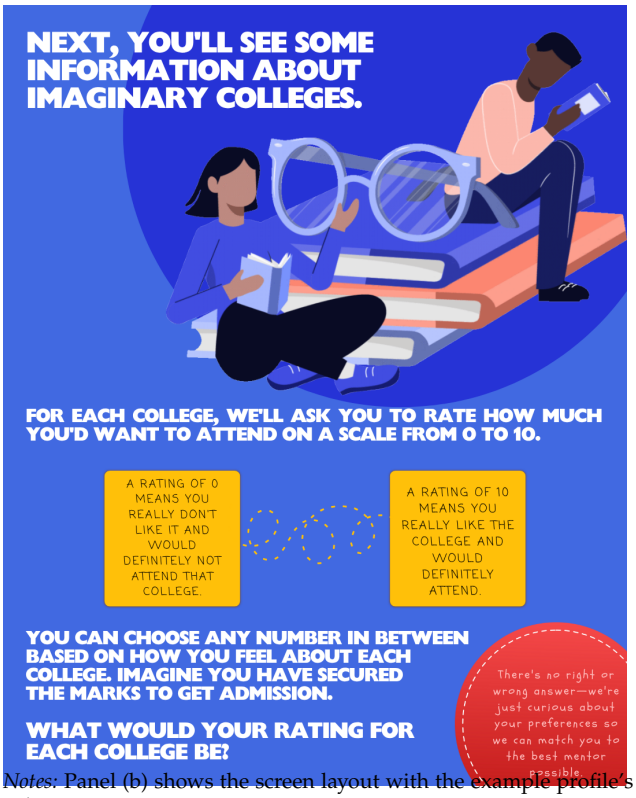
Notes: Monetary values were displayed in rupees. All randomized values were generated separately for each profile.

Presentation. Before the rating screens, students saw the note: "Please click on the "Play Button" to watch the video for instructions for the next section!" The same screen displayed an instruction poster and an instruction video. Each profile then appeared as a small table with the destination, yearly fees, monthly food and living costs, hostel information, admission cutoff, share selected for higher education, share placed in jobs, and average yearly earnings.

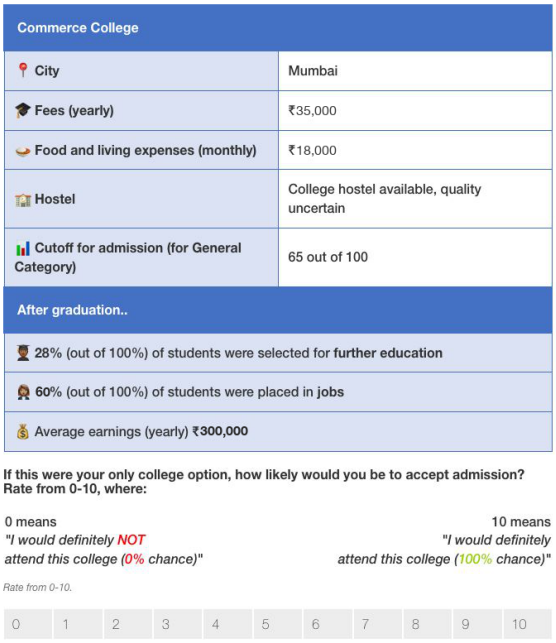
The rating prompt read: "If this were your only college option, how likely would you be to accept admission? Rate from 0-10, where:" The left endpoint defined 0 as "I would definitely NOT attend this college (0% chance)," and the right endpoint defined 10 as "I would definitely attend this college (100% chance)." The survey hint below the response scale read: "Rate from 0-10." Figure G1 shows the instruction poster and an example profile rating screen.

Figure G1: College rating exercise screens shown in the midline survey

(a) Instruction poster



(b) Profile rating screen



Notes: Panel (b) shows the screen layout with the example profile's fees, living costs, placement share and earnings set to typical values.

G.1 Willingness to pay to stay home

I recover a willingness to pay to remain at home from the baseline rating exercise with a hedonic regression of the 0–10 profile rating on an away indicator and the randomized profile attributes, using the same destination definition as the baseline discrete choice experiment: the home profile is the student's own district, displayed as the district-capital city, and the away profiles are Pune, Mumbai, or Nagpur. Ratings respond to fees but only weakly to earnings, so I express migration costs in fees, as the reduction in yearly fees that would make a student rate an away college as highly as the same college at home. I then express this amount as a share of the student's own expected earnings, so that it can be compared with the estimates from the choice experiment in Section 4.2. Table G2 reports the estimates overall and by gender.

Table G3 reports how much of the away rating gap the beliefs of Section 4.1 account for.

Table G2: Away rating gap and estimated migration costs in the baseline college-rating exercise

	All	Men	Women
Away rating gap (rating points)	-1.025***	-0.804***	-1.166***
	(0.119)	(0.188)	(0.154)
Yearly tuition (per 100 dollars)	-0.1117***	-0.0983***	-0.1175***
	(0.0174)	(0.0277)	(0.0224)
Monthly living cost (per 100 dollars)	-0.1123***	-0.1487***	-0.0905**
	(0.0338)	(0.0549)	(0.0430)
Placement share (per pp)	0.0063**	0.0035	0.0083**
	(0.0028)	(0.0045)	(0.0035)
Yearly earnings (per 100 dollars)	0.0027**	0.0050***	0.0011
	(0.0012)	(0.0018)	(0.0015)
Admission cutoff (per point)	-0.0052*	-0.0046	-0.0060
	(0.0029)	(0.0048)	(0.0037)
Higher-education share (per pp)	0.0076	0.0298**	-0.0063
	(0.0091)	(0.0140)	(0.0118)
Clean and safe hostel	0.427***	0.395	0.433**
	(0.153)	(0.246)	(0.194)
Hostel, quality uncertain	0.118	0.130	0.080
	(0.150)	(0.239)	(0.194)
Only private / PG hostel	0.076	0.058	0.086
	(0.153)	(0.239)	(0.197)
Migration costs (dollars of yearly tuition)	918***	818***	992***
	(184)	(296)	(242)
as % of median expected earnings	86	51	133
Median expected earnings (dollars a year)	1,067	1,600	747
Nr. students	1474	575	899
Nr. profiles	4422	1725	2697

Notes: OLS regression of the 0–10 rating of each college profile on an away indicator and the profile's randomized attributes, on the attentive sample of the choice experiment (1,474 students, three profiles each). Home is the student's own district, shown as the district capital, and away is Pune, Mumbai, or Nagpur. The away coefficient is the rating gap. Tuition, living costs, and earnings are per 100 dollars (93.75 INR per USD), placement and higher-education shares per percentage point, and the admission cutoff per point. The hostel indicators are relative to a profile with no hostel information. Ratings respond much more to fees than to earnings, so migration costs are estimated in tuition. The estimate is the away coefficient divided by the tuition coefficient, the yearly fee reduction that offsets studying away, in dollars and as a percentage of the group's median expected earnings. Men and women are estimated separately. Standard errors clustered by student in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table G3: Share of the away rating gap accounted for by beliefs

	All	Men	Women
Across profiles			
All eight beliefs	100.5	122.1	87.5
Leaving out moving difficulty	80.7	96.9	75.7
Within student (student fixed effects)			
All eight beliefs	85.2	84.3	89.4
Leaving out moving difficulty	65.6	56.7	72.2
Students	1,474	575	899

Notes: OLS regressions of the 0–10 rating of each college profile on an away indicator and the profile's randomized attributes, on the attentive sample of the choice experiment, without and with the student's beliefs about the profile's destination from the baseline survey. Each cell is the percentage by which the away coefficient falls once the beliefs are added. *Within student* adds student fixed effects. *All eight beliefs* uses the beliefs of Section 4.1; *Leaving out moving difficulty* drops that item, whose home value is set to "not challenging at all". Shares above 100 mean that the away coefficient changes sign. Men and women are estimated separately.

H Appendix: Belief and information questions

H.1 Non-pecuniary beliefs

At baseline, students answer the questions in Table H1 about each of the three destinations they are shown (Appendix E). In the wording, [destination] is the city being asked about. In the analysis every item is recoded so that higher values are better (safer, happier, more approving, easier to move, and so on).

Table H2 reports each item's baseline mean at home and away, and the away–home gap in home standard deviations, for all students and by gender.

Table H1: Baseline non-pecuniary belief questions

Belief	Question	Response scale
Happiness	If you live and study in [destination], how happy do you think you would be?	0–10 (0 = completely unhappy, 10 = completely happy)
Mental wellbeing	If you live and study in [destination], how likely is it that you might face difficulties with your mental wellbeing?	0–10 (0 = definitely would not, 10 = definitely would)
Safety	How safe would you feel living and studying in [destination]?	0–10 (0 = completely unsafe, 10 = completely safe)
Reputation back home	How likely is it that other people in [home] will talk badly about you, if you were living and studying in [destination]?	0–10 (0 = would never, 10 = would definitely)
Parental approval	All things considered, would your parents approve of you living and studying in [destination]?	0–10 (0 = would not approve at all, 10 = would completely approve)
Days seeing family	If you live and study in [destination], how many days per month would you meet family and friends in [home]?	Days per month, 0 to 30
Close friends	If you live and study in [destination], how many close friends do you think you would have in the first year of studying there?	0 to 10, or more than 10 (coded as 11)
Moving difficulty	Imagine you are moving to [destination] for college (finding accommodation, registering, locating healthcare and food, learning to navigate the city and transport). How challenging would it be to move to [destination] for college?	Four points: not challenging at all, slightly, very, overwhelmingly challenging

Table H2: Baseline beliefs about destinations: home versus away

			Away − home (home SDs)		
	Home mean	Away mean	All	Men	Women
Emotional well-being					
Happiness	7.37	5.75	-0.53***	-0.54***	-0.53***
			(0.028)	(0.043)	(0.037)
Mental wellbeing	0.63	0.55	-0.22***	-0.15***	-0.26***
			(0.025)	(0.039)	(0.032)
Social ties					
Reputation back home	0.66	0.61	-0.14***	-0.04	-0.20***
			(0.022)	(0.034)	(0.029)
Days seeing family	9.04	7.32	-0.18***	-0.19***	-0.18***
			(0.019)	(0.031)	(0.023)
Close friends at destination	3.67	2.67	-0.28***	-0.27***	-0.29***
			(0.019)	(0.032)	(0.024)
Parental approval	6.98	4.96	-0.62***	-0.46***	-0.72***
			(0.026)	(0.041)	(0.033)
Practical and lived experience					
Safety	7.43	5.28	-0.72***	-0.60***	-0.79***
			(0.028)	(0.046)	(0.036)
Moving difficulty[a]	3.00	1.59	-1.67***	-1.50***	-1.79***
			(0.020)	(0.033)	(0.026)

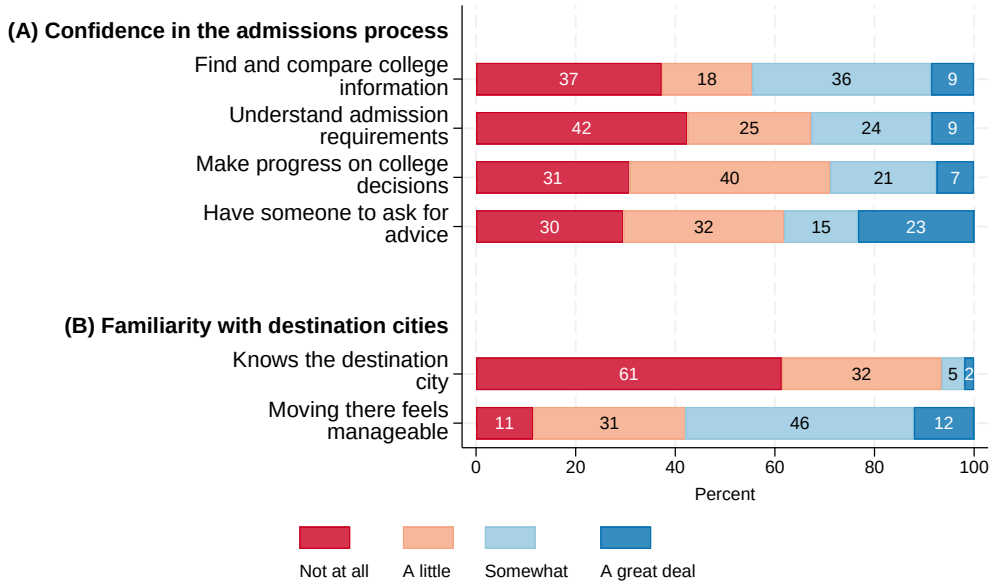
Notes: Home and away means in each item's own units (0–10 scales, probabilities, or counts). The gap columns are coefficients from regressions of the item, standardized by its home mean and standard deviation, on an away indicator, pooling all away destinations, for all students and by gender. At baseline every away destination lies outside the home district. Standard errors clustered by student. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[a] Moving difficulty is asked only about away destinations. Its home value is set to "not challenging at all", and it is standardized by its standard deviation across away destinations.

H.2 Information items

Figure H1 reports the distribution of the baseline items on navigating college admissions and on familiarity with the destination cities that underlie Fact 3 (Section 4.4).

Figure H1: Baseline knowledge about college admissions and about destination cities



Notes: Each bar shows the distribution of answers to one baseline item, from *Not at all* to *A great deal*. Panel A asks how difficult it is for the student to find and compare college information, understand admission requirements, make progress on college decisions, and find someone to ask for advice; *Not at all* is "very difficult, I don't know how to do it" and *A great deal* is "not difficult at all, I know how to do it". Panel B pools the destination cities each student was asked about and shows how much the student knows about studying and student life there (*Knows the destination city*) and the reverse of how challenging moving there would be (*Moving there feels manageable*).

I Appendix: How the program was presented

This appendix describes the steps through which students learned about Campus Trail and joined the study, and what they were told at each step. The sampling of schools and the randomization are described in Section 5.

School visit During each school visit, an enumerator introduces Campus Trail to all final-year students present. Enumerators are given the following script for the introduction.

> Good morning. My name is [name]. We are here today to introduce you to a program called Campus Trail, which is a collaboration between researchers at Stockholm University in Sweden and DAI Research & Advisory Services. The program is a very

helpful resource for students like you who are at the beginning of their education and need to prepare for college.

First, I would like to know what your plans are after you complete your 12th board exams. Are you preparing for JEE, NEET, or CET?[22] How many of you have already started to look into which colleges you should apply to or which degrees you should look for? Have you looked up the cutoff scores for different colleges, or have you started applying for the entrance exams for these colleges? Have you thought about how difficult it could be to find a place to stay when you get into a college, or what hostels are like? Finding friends and people to spend time with? Or even finding good places to eat when you move to a new city?

All of this can seem overwhelming, and it can be difficult for a student like you to know about all the processes for getting into your dream colleges. You must feel the pressure to score well, but you should also be prepared to go through these processes after your board exams.

This is where the Campus Trail program helps you. Think about having an older brother or sister, or an older neighbor, who already went to college, for example to COEP in Pune or to a government medical college. They can help you understand what college life will be like and how to get into these colleges. Campus Trail will help you in a similar way. We will connect you with current students at some of the colleges you are thinking of going to. These students are similar to you, and they have migrated to big cities to study there. These students will be your mentors. They are not much older than you and come from smaller cities like Ahmednagar (now Ahilyanagar), Sangli, or Bhandara, but they have been through the admission process, found places to live, managed to get scholarships, and settled into these colleges.

You will get to speak to these students one-on-one and discuss your problems or doubts privately, and they can help you choose a good educational option.

Before that, you will need to fill in a survey, which we will now give you. The survey asks about yourself, your goals and what you dream of doing, where you would like to study, and how you make decisions. It takes about 45 minutes to complete, and it will help us connect you to the mentor who can best help you with your needs. Once you complete the survey, you will enter a lottery, in which some students will be randomly chosen to connect with a mentor. Do not think of this as a test in which your answers decide whether you get a mentor. No answer is right or wrong. We just want to know about your goals and choices.

[22] JEE and NEET are the national entrance exams for engineering and for medicine. CET is the state common entrance test for engineering, pharmacy, and management. COEP, mentioned below, is the College of Engineering Pune, a public engineering college.

> You can choose not to take part in the survey, or not to take part in the mentorship sessions, at any point in the program. However, we believe a student like you will benefit greatly from this program. The program is free, and you do not need to pay or spend any money.
>
> If you have any questions about the program or the survey, you can ask them now.
>
> Could all students who are really interested in participating, and who will complete the survey carefully and seriously, please write their names on the sign-up sheet?

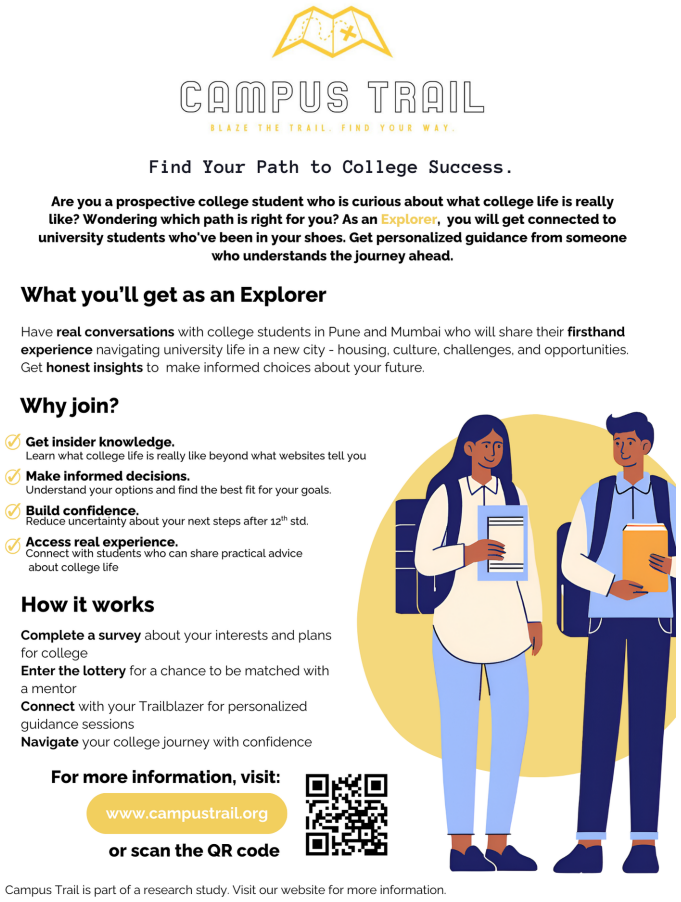
Sign-up Students who want to take part write their names on the sign-up sheet. Only these students take the baseline survey and enter the study.

Baseline survey Students who sign up fill in the baseline survey on a tablet in school. The survey opens with two short videos. The first is a three-minute animation, narrated in Marathi, that explains what the program offers and how the survey answers are used. In it, current college students, called Trailblazers, share their experience of college, including academics, hostel life, and career opportunities. The video explains that the survey is used to match each student with a suitable Trailblazer and shows a student who enters their interests and goals and is then matched with a mentor. It asks students to answer honestly, stresses that there are no right or wrong answers, and tells students who have not yet decided what to study that the program is for them as well. The second video familiarizes students with the question types used in the survey, such as choosing one option, ticking all options that apply, choosing from a drop-down list, and choosing between two alternatives, the format of the choice experiment. Further instruction videos appear later in the survey, before the choice experiment and the college rating exercise.

Flyer for parents Students who sign up also receive a printed flyer and a pen as a small thank-you, and are asked to take the flyer home to their parents. The flyer informs parents about the program and gives students an accurate written description of it (Figure I1). It calls participating students Explorers and mentors Trailblazers. It describes the program as conversations with college students in Pune and Mumbai about university life in a new city, including housing, culture, challenges, and opportunities. It lists three steps, a survey about the student's interests and plans for college, a lottery for the chance to be matched with a mentor, and guidance sessions with the mentor. It also states that Campus Trail is part of a research study and refers to the program website (Figure I2). The website describes the program in the same terms. It adds an

example of a twelfth-grade student who is considering Pune or Mumbai and is matched with a student who moved to Pune two years earlier, and it states that students can join only through their school, that taking part in the research is voluntary, and that responses are kept confidential. Most of the website is addressed to prospective mentors (Appendix D).

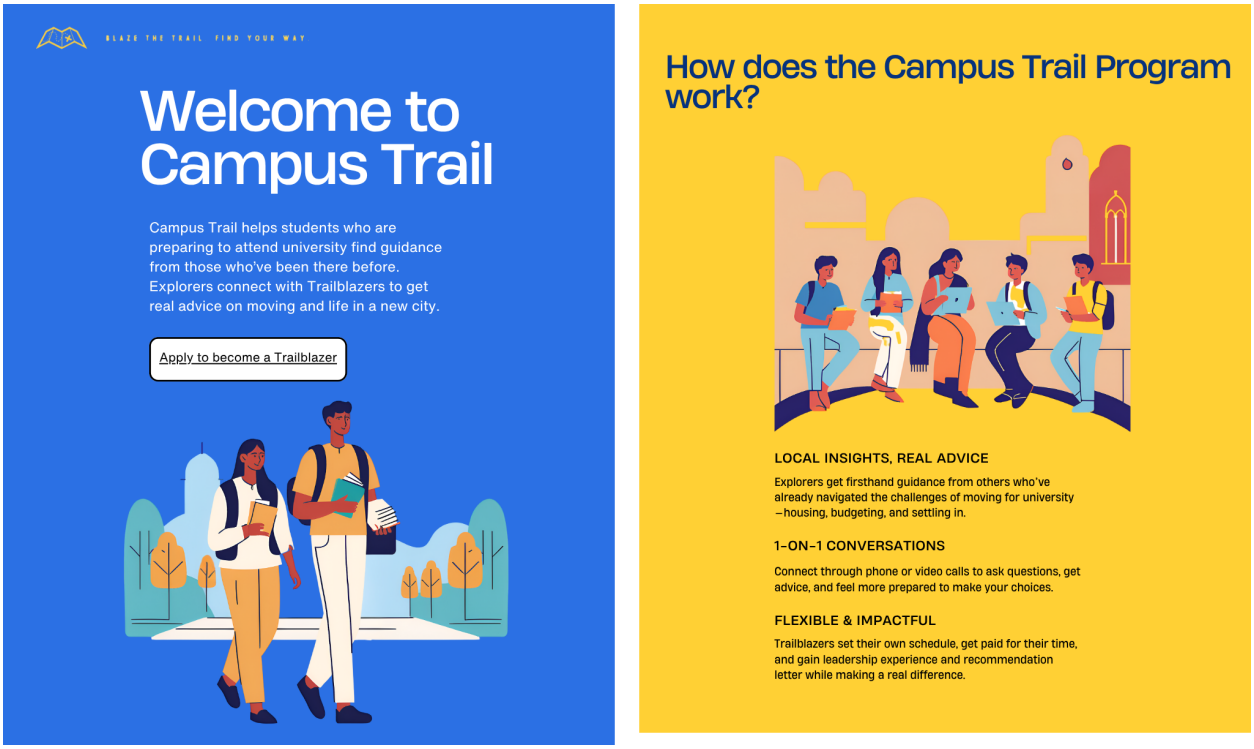
Figure I1: Program flyer given to students



Notes: The flyer was printed and given to students who signed up, to take home to their parents.

Lottery and matching After the baseline survey, students are randomly assigned to the program or to the control group, and each student in the program is matched with a mentor using their survey answers (Section 5). Students are told about both steps when they sign up. Their answers therefore affect which mentor they get, not whether they get one.

Figure I2: Program website



Notes: The left panel shows the top of the program website (campustrail.org), and the right panel shows the section that explains how the program works. Captured on 22 October 2026. The rest of the website, including the mentor profiles, is not shown.

J Appendix: Endline data collection and attrition

This appendix details the endline survey implementation, the tracking protocol used to limit attrition, and the associated weighting and bounding procedures summarized in Section 5.

J.1 Survey implementation

Endline surveys are conducted over the phone and take on average twenty minutes; respondents are offered 150 INR (about 1.6 USD) for completing the interview. Whenever a student is unavailable, the field team conducts a proxy interview with a parent or another family member, collecting information on the primary outcomes only. Proxy interviews take on average five minutes and are not incentivized. The main analysis uses students' own answers and, when only a proxy interview is available, proxy-reported outcomes; Appendix Table L2 excludes proxy interviews. Where both are available, the student's answer is used.

J.2 Tracking and weighting

Endline data collection begins with a standard phone-tracking phase in which enumerators attempt to reach every student using all available numbers (the student's own and those of parents or other household members) with repeated attempts at different times of day and attempts to obtain updated contact information from parents and schools.

After this phase I construct the intensive-tracking frame. A student is classified as "not reached" if, by a pre-specified tracking checkpoint, no student or proxy interview has been completed. The intensive-tracking unit is the school, the natural cluster for tracing students: students were recruited through their schools, and schools hold updated contact information through records, teachers, and classmates. Letting N_s denote the number of not-reached students in school s at the checkpoint, schools with at least five not-reached students are eligible for intensive tracking.[23] I select up to ten schools using sequential probability-proportional-to-size sampling without replacement, drawing each school with probability proportional to its number of not-reached students among schools not yet selected. In selected schools the field team traces all not-reached students by phone through their classmates, who are asked for updated contact details, and completes student interviews where possible and proxy interviews otherwise. These students are offered 250 INR, in place of 150 INR, for completing the interview.

For each school I compute the inclusion probability π_s implied by this procedure by simulation, conditional on the realized N_s. I report unweighted estimates using all completed surveys and weighted estimates that account for the unequal probability of intensive tracking, with weights

$$w_i = \begin{cases} 1, & \text{if } i \text{ was surveyed before the tracking checkpoint,} \\ 1/\pi_s, & \text{if } i \text{ was surveyed through intensive tracking.} \end{cases}$$

J.3 Response and attrition

I report overall and arm-specific attrition at the checkpoint and after intensive tracking, test for differential attrition by treatment arm, and test whether baseline covariates predict attrition and do so differentially by treatment. Table J1 reports response by arm and wave under the sample definitions used in the paper. Table J2 reports endline response by treatment arm on the full experimental sample, and Table J3 reports, for both survey waves, whether baseline characteristics predict response (selective attrition) and whether that selection differs by arm (differential attrition). The midline was adminis-

[23] I exclude the smallest schools (those with fewer than five not-reached students) up to at most 5% of the attrited sample.

tered in person during school visits and involved no tracking: enumerators surveyed the students who were present, so midline non-response is selective on who attended school that day but does not differ by treatment arm. The endline, reached by phone with the protocol above, shows mild differential attrition concentrated among treated students who intended to stay home at baseline.

Table J1: Survey response and attrition, by treatment arm

	Control	Treated	Difference
Panel A. Midline (in school, untracked)			
Surveyed	0.717	0.714	-0.003
			(0.022)
Panel B. Endline (by phone, tracked)			
Any interview	0.848	0.838	-0.010
			(0.018)
Direct student interview	0.674	0.685	0.011
			(0.023)
Proxy interview only	0.174	0.153	-0.022
			(0.018)
Full primary data (student)	0.686	0.708	0.022
			(0.022)
Full primary data (any source)	0.791	0.797	0.006
			(0.020)
Observations		1,682	

Notes: Share of the full experimental sample responding at each wave, by treatment arm, with the treatment–control difference from an OLS regression and its robust standard error below. Panel A reports the midline response rate; the midline was administered in schools without tracking, so response reflects who attended on the survey day. Panel B reports endline response under alternative sample definitions: *Any interview* (a completed student or proxy interview), *Direct student interview*, *Proxy interview only*, and *Full primary data* observed from the student or from any source. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table J2: Endline response and estimates under intensive-tracking weights

Panel A. Survey response, by treatment arm

	Control	Treated	Difference
Any interview at the checkpoint	0.757	0.735	-0.022
			(0.021)
Any interview	0.848	0.838	-0.010
			(0.018)
Student interview	0.674	0.685	0.011
			(0.023)
Proxy interview	0.174	0.153	-0.022
			(0.018)
Full primary (student)	0.686	0.708	0.022
			(0.022)
Full primary (any)	0.791	0.797	0.006
			(0.020)
Recovered through intensive tracking		92 of 109 pursued	
Effective response rate		0.908	
Observations		1682	

Panel B. Primary outcomes, with and without intensive-tracking weights

	Control mean	Unweighted	Weighted
Out-of-district index	-0.000	0.097^{*}	0.076
		(0.053)	(0.056)
Migrated out of district	0.120	-0.002	0.006
		(0.017)	(0.017)
Away-from-home index	0.000	0.155^{***}	0.125^{**}
		(0.054)	(0.056)
Migrated away from home	0.216	0.027	0.035
		(0.022)	(0.022)

Notes: Panel A is the full experimental sample. Each row reports, for each arm, the share of students meeting the row definition; the Difference column reports the OLS coefficient on treatment assignment, with the robust standard error in parentheses below. *Any interview at the checkpoint*: a completed student or proxy interview by the tracking checkpoint, the point at which the not-reached frame was fixed and the intensive-tracking schools drawn. *Any interview*: a completed student or proxy interview at the end of fieldwork, inclusive of students recovered since the checkpoint. *Student interview*: a completed interview with the student directly. *Proxy interview*: a completed proxy interview (a parent or another family member) with no completed student interview. *Full primary (student)*: all six primary-outcome components observed from the student's own report, including partially completed interviews. *Full primary (any)*: all six components observed from any source. The recovery line reports how many of the not-reached students pursued in the drawn schools were interviewed. *Effective response rate*: students interviewed by the checkpoint plus students recovered through intensive tracking, each weighted by the inverse probability that their school was drawn, as a share of the full sample.

Panel B reports the primary outcomes on students with a completed interview, first unweighted and then weighted by the pre-specified intensive-tracking weight defined above: one for students surveyed before the tracking checkpoint, and $1/\pi_s$ for students recovered through intensive tracking, where π_s is the simulated probability that their school was drawn. A recovered student therefore stands in for the comparable not-reached students in the eligible schools that were not drawn. The two indices are inverse-covariance-weighted (Anderson, 2008). Not-reached students in schools ineligible for intensive tracking have no probability of recovery and are represented by neither column; they are assigned worst-case values in the bounds reported below. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table J3: Selective and differential attrition at midline and endline

	Midline (in school, untracked)		Endline (phone, tracked)	
	Selective (X)	Differential $(T \times X)$	Selective (X)	Differential $(T \times X)$
Baseline demographics				
Female	0.104***	0.008	0.029	-0.014
	(0.023)	(0.046)	(0.018)	(0.037)
Past board marks (SD)	0.029**	0.016	0.016*	0.000
	(0.011)	(0.023)	(0.009)	(0.018)
Self-efficacy (SD)	0.018	-0.023	0.028***	-0.037*
	(0.011)	(0.022)	(0.009)	(0.019)
Asset index (SD)	-0.016	0.039*	0.003	0.008
	(0.012)	(0.023)	(0.008)	(0.017)
Household size (SD)	0.023**	-0.029	0.005	-0.019
	(0.011)	(0.022)	(0.009)	(0.019)
Internet access	-0.042	-0.033	-0.004	-0.064
	(0.028)	(0.056)	(0.023)	(0.046)
Mother class 10	0.004	-0.012	0.025	0.033
	(0.022)	(0.044)	(0.018)	(0.036)
Father class 10	-0.020	-0.066	0.024	0.036
	(0.024)	(0.047)	(0.020)	(0.039)
Science stream	0.015	0.043	0.016	0.071*
	(0.024)	(0.048)	(0.019)	(0.039)
General social category	-0.009	0.087*	0.021	0.057
	(0.023)	(0.046)	(0.018)	(0.036)
Sangli district	-0.030	0.048	-0.053***	0.029
	(0.022)	(0.044)	(0.018)	(0.036)
Lives in a district capital	-0.042	-0.006	0.005	0.009
	(0.026)	(0.052)	(0.020)	(0.041)
Baseline analogues of primary outcomes				
Wants to live elsewhere at 25	0.002	-0.017	0.035*	0.091**
	(0.022)	(0.044)	(0.018)	(0.036)
P(apply) away from home (SD)	0.009	0.021	0.013	0.008
	(0.011)	(0.022)	(0.009)	(0.018)
Rating of away destinations (SD)	0.010	0.014	0.008	-0.031*
	(0.011)	(0.022)	(0.009)	(0.018)
Any network away from home	-0.005	-0.038	0.019	0.045
	(0.029)	(0.058)	(0.023)	(0.047)
Joint test p	0.001	0.362	0.107	0.010
Observations		1682		

Notes: Each wave's two columns report OLS regressions of a response indicator on baseline characteristics, estimated on the full 1,682-student experimental sample. The *Selective* column regresses the indicator for being surveyed on the baseline characteristic alone; its coefficient shows whether the characteristic predicts response. The *Differential* column regresses response on the characteristic, treatment assignment, and their interaction, and reports the interaction coefficient; it shows whether selection into response differs by arm. Continuous predictors are standardized to control-group units, so their coefficients are per standard deviation; binary predictors are in levels. Midline response is a completed midline interview; endline response is a completed student or proxy interview. The midline was administered in person during school visits without tracking, so midline response reflects who attended school on the survey day; the endline was conducted by phone under the tracking protocol described above. "Joint test p" is the p-value of an F-test that all baseline characteristics jointly predict response (*Selective* columns) or that all treatment interactions are jointly zero (*Differential* columns). Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J.4 Attrition bounds

I report three sets of bounds following Coppock et al. (2017). First, for the binary primary outcomes I report double-sampling worst-case bounds, using the randomized intensive-tracking sample to estimate outcomes among initially not-reached students and assigning worst-case values to cases that remain unrecovered. Second, I report sensitivity bounds in which the never-surveyed are assumed to differ from the recovered by a range of magnitudes calibrated to the observed gap between students reached during standard phone tracking and students reached through intensive tracking. Third, I report post-stratified double-sampling bounds, defining poststrata by baseline gender and stream and aggregating poststratum-specific treatment–control differences using the baseline share of each poststratum, collapsing to gender only if the gender-by-stream cells are too sparse. Students in schools ineligible for intensive tracking have zero recovery probability and are assigned worst-case values.

Table J4 reports these procedures for the migration index. Intensive-tracking weights lower the point estimate for any move away from home from 0.15 to 0.12 standard deviations. The double-sampling worst-case and poststratified bounds are uninformative, because every student who remains unrecovered, 16% of the sample, is assigned a worst-case value. For the lower bound, unrecovered treated students receive the lowest value of the index observed in the sample, as if they had not applied to or enrolled in any further study, and unrecovered control students the highest. With about 90% of students applying for further study, this is unrealistic. Many unrecovered students were reached but refused the interview. Their non-response is a refusal of the survey rather than a move away from home, which is the main concern with attrition in this setting, and for some treated students among them, the mentor reports that they plan to continue studying at home. I therefore also report worst-case bounds that assign the worst case only to the students who could not be found, excluding the 190 who explicitly refused the interview: the interval is then 0.08 to 0.19. The sensitivity bounds, which shift the never-surveyed by up to twice the observed tracking gap, span 0.10 to 0.15, and assigning each unrecovered student the control-group prediction at their own baseline characteristics, which assumes the program had no effect on them, gives 0.11 (bootstrap standard error 0.05).

Table J4: Attrition adjustments and bounds for the migration index

	Out of district	Away from home
Point estimates		
Complete case	0.097*	0.155***
	(0.053)	(0.054)
Intensive-tracking weighted	0.076	0.125**
	(0.056)	(0.056)
Unrecovered at control prediction[a]	0.068	0.109**
	(0.051)	(0.050)
Double-sampling worst case		
Full frame	[-0.484, 0.619]	[-0.382, 0.611]
Poststratified (gender × stream)	[-0.485, 0.610]	[-0.386, 0.600]
Poststratified (gender only)	[-0.485, 0.618]	[-0.384, 0.609]
Refusals excluded (190 students)[a]	[0.018, 0.135]	[0.081, 0.186]
Sensitivity, missing shifted by the tracking gap		
calibration gap g	0.009	-0.053
both arms at $+g$ (point)	0.076	0.125
both arms at $+2g$ (point)	0.076	0.126
adversarial, $\pm g$	[0.074, 0.079]	[0.111, 0.138]
adversarial, $\pm 2g$	[0.072, 0.081]	[0.098, 0.151]

Notes: Intention-to-treat estimates for the pre-specified migration index under the attrition procedures of the pre-analysis plan, for moves out of the home district (first column) and any move away from home (second column). *Complete case*: all students with a completed interview. *Intensive-tracking weighted*: weights of one for students surveyed before the tracking checkpoint and $1/\pi_s$ for students recovered through intensive tracking. *Unrecovered at control prediction*: each unrecovered student is assigned the control-group prediction at their own baseline characteristics. Robust standard errors in parentheses, and for the last point estimate a bootstrap standard error over the imputation and estimation. *Double-sampling worst case*: bounds that assign the worst-case value of the index to students who remain unrecovered, in the full frame, within gender×stream or gender poststrata aggregated with baseline shares, or excluding students who explicitly refused the interview. *Sensitivity*: the never-surveyed are assumed to differ from the recovered by the calibration gap g, the observed difference between students reached in standard tracking and students reached through intensive tracking, in the same direction in both arms (point) or adversarially across arms (intervals), at one and two times g. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
[a] Not pre-specified; Appendix J.4 explains why.

K Appendix: Parent survey and parent–student comparisons

This appendix describes the parent survey, who was reached, and how parents' preferences and beliefs about studying away compare with their children's.

Data collection. The parent instrument mirrors the student baseline (Section 2.2): educational aspirations and expectations for the child, affordability of tuition and monthly expenses, a discrete choice experiment over hypothetical college profiles (four choice sets, same three numeraire arms as the student DCE), ratings of the three destinations

in the student's elicitation set, a two-city belief module covering home and the nearer of the student's two elicited cities (city order randomized), networks, destination consideration, and the gender-attitude and family-decision batteries. Parents were more likely to be reached when the student was a woman, came from a less wealthy household, or reported higher self-efficacy or a higher likelihood of applying away from home: baseline covariates jointly predict which parents were reached ($p = 0.003$), though this selection does not differ by treatment arm ($p = 0.94$; Table K1).

Parents' migration costs are a multiple of their children's. Estimating the same conditional logit as in Section 4.1 on the parent DCE, the average parent would forgo roughly a 450 % increase in the child's post-graduation earnings to keep them at home, versus roughly 140 % for students when the student experiment is estimated with the same specification on the full baseline sample (the pooled-wave estimate of Section 4.2 is 129%; employment arm: 168 pp versus 31 pp). In the cutoff arm the comparison is made in coefficient units: the parent away coefficient is three times the student one (-2.28 vs. -0.74), but parents barely respond to the cutoff attribute, so no meaningful rank-denominated MRS exists for them; parent-arm estimates are in any case imprecise, with ≈270 parents and four choice sets per arm. The destination ratings tell the same story in simpler form: the rating gap for away destinations is -5.4 points (0–10) for parents against -1.5 for students, driven both by parents rating home higher (9.2 vs. 7.6) and away cities lower (3.8 vs. 6.1).

Parents perceive a larger gap on every non-pecuniary belief. Parents expect away destinations to compare worse to home than students do on each elicited item: expected happiness falls by 4.1 points (0–10) for parents versus 1.2 for students, safety by 4.7 versus 2.0, days seeing family by 9.6 versus 1.5 per month, and parental approval by 4.7 versus 1.9.

Students misperceive their parents' main objection. Both sides are asked the *main* reason the parent would object to the child attending college in the distant city, from the same reason list. Only 13.8 % of matched pairs name the same reason. The disagreement is systematic: parents overwhelmingly name safety (34.6 % of parents, versus 7.8 % of students guessing safety), while students over-attribute concerns about housing (14.3 % vs. 5.9 %) and jobs (10.2 % vs. 0.9 %). Both sides agree on expense (26.1 % vs. 22.5 %).

Table K1: Selection into the parent survey

	Predicts response (selective)	Differential by arm (Treated × X)
Female	0.048*	-0.069
	(0.025)	(0.050)
Age	-0.004	-0.003
	(0.012)	(0.024)
General category	0.005	-0.007
	(0.025)	(0.051)
Science stream	-0.002	-0.010
	(0.026)	(0.052)
Past board marks	0.011	-0.021
	(0.012)	(0.024)
Asset index	-0.031**	0.035
	(0.012)	(0.024)
Household size	-0.015	-0.016
	(0.012)	(0.024)
Mother completed class 10	-0.038	-0.041
	(0.024)	(0.049)
Father completed class 10	-0.005	0.002
	(0.026)	(0.053)
Internet access	-0.047	-0.028
	(0.032)	(0.064)
Self-efficacy	0.030**	-0.006
	(0.012)	(0.024)
Any network away from home	-0.020	-0.006
	(0.025)	(0.049)
P(apply) away from home	0.031**	-0.010
	(0.012)	(0.024)
Joint test p	0.003	0.938
Observations	1,682	

Notes: Selection into the parent survey on the full experimental sample ($N = 1{,}682$). The *Predicts response* column reports, for each baseline characteristic, the coefficient from a regression of an indicator for the student's parent being surveyed on that characteristic (selective attrition); the *Differential by arm* column reports the coefficient on the characteristic's interaction with treatment assignment (whether that selection differs by arm). Continuous predictors are standardized, so their coefficients are per standard deviation; binary predictors are in levels. Baseline characteristics jointly predict which parents were reached ($p = 0.003$), but this selection does not differ by treatment arm ($p = 0.94$). Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Parental attitudes and program engagement. Parents' willingness for the child to migrate (listing any away destination when asked where the child should look for colleges: 44 %; approval of the child living in the nearer elicited city: 4.3 on a 0–10 scale) predicts mentorship take-up only weakly on its own and adds nothing once the student's own baseline application intentions are controlled for (parent measures joint $p = 0.75$): engagement with the mentor is driven by the student's own willingness.

L Appendix: Robustness

This appendix checks the main estimates against alternative definitions and samples, and tests for spillovers to control students.

L.1 Alternative definitions and samples

Table L1 re-estimates the migration index under alternative constructions and samples: the pre-specified index on all students with a completed interview; the same index with post-double-selection lasso controls; direct student interviews only; an expanded sample that adds resolved cases missing some components together with incomplete surveys that reached the primary module; and an index built from realized outcomes only, dropping planned actions from every component. Every index is re-standardized within its own estimation sample, so control means are zero by construction and the columns are directly comparable in control standard deviations. The effect is present in every column, between 0.13 and 0.20 standard deviations; excluding proxy interviews lowers it slightly and counting only realized outcomes raises it.

Table L2 reports every outcome on direct student interviews only. Table L3 defines moves as leaving the home district rather than leaving home. The index rises by 0.10 standard deviations, through applying and enrolling, which do not depend on the destination; applications and moves out of the home district do not change.

Table L1: Migration index under alternative definitions and samples

	Pre-specified	PDS controls	Direct interviews	Incl. incomplete	Realized outcomes only
Assigned to mentorship	0.155***	0.159***	0.141**	0.133***	0.197***
	(0.054)	(0.053)	(0.058)	(0.050)	(0.065)
Control mean	0.000	0.000	0.000	-0.000	0.000
Observations	1299	1299	1115	1485	1110

Notes: Each cell is the intention-to-treat coefficient on assignment to mentorship for the inverse-covariance-weighted migration index (Anderson, 2008), any move away from home, under one construction, with the standard error below. *Pre-specified*: the six-component index on all students with a completed interview. *PDS controls*: adds post-double-selection lasso controls chosen from the baseline covariate set, with analytic standard errors. *Direct interviews*: restricts to direct student interviews. *Incl. incomplete*: the expanded index, adding resolved cases missing some components together with incomplete surveys that reached the primary module. *Realized outcomes only*: drops planned actions from every component and counts only applications made, financing applied for, enrollment started, and moves completed by the interview date, on direct student interviews. Each index is re-standardized within its own estimation sample, so control means are zero by construction. Robust standard errors except in the PDS column. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table L2: Effects on educational migration, direct student interviews only

	Index of educational migration	Applied to study	Applied to education financing	Enrolled in a program	Applied to study away	Migrated away	Migrated for education
Assigned to mentorship	0.141**	0.024*	0.017	0.033	0.047	0.036	0.024
	(0.058)	(0.013)	(0.029)	(0.023)	(0.029)	(0.025)	(0.024)
Control mean	0.000	0.937	0.579	0.809	0.346	0.212	0.181
Observations	1115	1133	1136	1123	1128	1140	1123

Notes: Each column reports the coefficient from an OLS regression of the outcome on an indicator for assignment to the mentorship program. Robust standard errors in parentheses. The sample is all students with a completed direct endline interview; proxy interviews are excluded, and the index is re-estimated within this sample. All outcomes except the index are indicators equal to one if the student has taken the action or plans to within two months. Away from home means outside the student's home sub-district. *Applied to study*: applied for further study. *Applied to education financing*: applied for a scholarship or loan. *Enrolled in a program*: enrolled in a degree college, vocational institute, or coaching program. *Applied to study away*: applied to study away from home. *Migrated away*: moved away from home. *Migrated for education*: migrated away and enrolled. *Index of educational migration*: inverse-covariance-weighted average of the six components (Anderson, 2008), for students with all six observed, standardized to mean zero and standard deviation one in the control group. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table L3: Effects on out-of-district educational migration

	Index of educational migration	Applied to study	Applied to education financing	Enrolled in a program	Applied to study away	Migrated away	Migrated for education
Panel A. All students (student or proxy interview)							
Assigned to mentorship	0.097*	0.031**	0.029	0.046**	0.002	-0.002	-0.003
	(0.053)	(0.014)	(0.027)	(0.021)	(0.021)	(0.017)	(0.016)
Control mean	-0.000	0.906	0.521	0.788	0.173	0.120	0.098
Observations	1299	1391	1357	1383	1365	1403	1374
Panel B. Direct student interviews							
Assigned to mentorship	0.075	0.024*	0.017	0.033	0.003	-0.003	-0.001
	(0.058)	(0.013)	(0.029)	(0.023)	(0.023)	(0.019)	(0.018)
Control mean	0.000	0.937	0.579	0.809	0.182	0.122	0.099
Observations	1115	1133	1136	1123	1128	1140	1123

Notes: Each column reports the coefficient from an OLS regression of the outcome on an indicator for assignment to the mentorship program. Robust standard errors in parentheses. Panel A uses all students with a completed endline interview; for students who could not be interviewed, outcomes are reported by a parent or another family member. Panel B uses direct student interviews only, with the index re-estimated within that sample. All outcomes except the index are indicators equal to one if the student has taken the action or plans to within two months. In this table, away means outside the student's home district. *Applied to study*: applied for further study. *Applied to education financing*: applied for a scholarship or loan. *Enrolled in a program*: enrolled in a degree college, vocational institute, or coaching program. *Applied to study away*: applied to study outside the home district. *Migrated away*: moved outside the home district. *Migrated for education*: moved outside the home district and enrolled. *Index of educational migration*: inverse-covariance-weighted average of the six components (Anderson, 2008), for students with all six observed, standardized to mean zero and standard deviation one in the control group. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

L.2 Spillovers between students

Because randomization was at the individual level within schools, treated students could in principle relay information to control classmates, which would attenuate the intention-to-treat estimates. I test for such interference using the best-friend nominations collected at midline, which identify, for most students, a specific classmate in the experimental sample. Control students whose best friend was assigned to treatment are no more likely to migrate or to apply to colleges away from home: the coefficient on having a treated best friend is -0.04 (s.e. 0.05) for any move away from home, and the estimates for the other primary outcomes are similarly small and, if anything, negative. The result is unchanged when I restrict to friendships of at least a year's standing, which predate the intervention and so cannot reflect any post-treatment change in whom a student names as a best friend (-0.04, s.e. 0.06). I find no evidence that spillovers to control students attenuate the estimates.

M Appendix: Heterogeneity and other outcomes

This appendix reports the pre-specified heterogeneity analyses and the remaining pre-specified outcomes.

M.1 Heterogeneity

Tables M1 and M2 split the main estimates by academic stream and by whether the student knew anyone away from home at baseline, and Table M3 by decision-making power, separately for men and women. The effects are concentrated among commerce and arts students; they do not differ by baseline contacts or, within gender, by decision-making power.

Table M1: Effects on educational migration, by academic stream

	Index of educational migration	Applied to study	Applied to education financing	Enrolled in a program	Applied to study away	Migrated away	Migrated for education
Panel A. Science							
Assigned to mentorship	0.072	0.013	0.001	0.040	-0.000	-0.008	-0.017
	(0.064)	(0.017)	(0.032)	(0.026)	(0.032)	(0.029)	(0.027)
Control mean	0.162	0.922	0.577	0.774	0.439	0.282	0.234
Observations	894	953	930	949	943	959	941
Panel B. Commerce or arts							
Assigned to mentorship	0.286***	0.068**	0.072	0.065*	0.095***	0.087***	0.086***
	(0.093)	(0.027)	(0.048)	(0.034)	(0.037)	(0.032)	(0.031)
Control mean	-0.325	0.873	0.410	0.815	0.122	0.084	0.073
Observations	405	438	427	434	422	444	433
p-value, commerce or arts = science	0.058	0.088	0.224	0.567	0.051	0.027	0.012

Notes: Each column reports estimates from an OLS regression of the outcome on an indicator for assignment to the mentorship program, an indicator for the commerce or arts stream, and their interaction. Robust standard errors in parentheses. Panel A reports the coefficient on assignment (science). Panel B reports the sum of the coefficients on assignment and the interaction (commerce or arts). The last row reports the *p*-value of the interaction coefficient. Control means are for the control group of each stream. The sample is all students with a completed endline interview; for students who could not be interviewed, outcomes are reported by a parent or another family member. All outcomes except the index are indicators equal to one if the student has taken the action or plans to within two months. Away from home means outside the student's home sub-district. *Applied to study*: applied for further study. *Applied to education financing*: applied for a scholarship or loan. *Enrolled in a program*: enrolled in a degree college, vocational institute, or coaching program. *Applied to study away*: applied to study away from home. *Migrated away*: moved away from home. *Migrated for education*: migrated away and enrolled. *Index of educational migration*: inverse-covariance-weighted average of the six components (Anderson, 2008), for students with all six observed, standardized to mean zero and standard deviation one in the control group. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table M2: Effects on educational migration, by baseline contacts away from home

	Index of educational migration	Applied to study	Applied to education financing	Enrolled in a program	Applied to study away	Migrated away	Migrated for education
Panel A. Knows no one away							
Assigned to mentorship	0.115	0.034	-0.030	0.049	0.040	0.029	0.019
	(0.084)	(0.023)	(0.042)	(0.032)	(0.039)	(0.034)	(0.032)
Control mean	-0.008	0.899	0.575	0.790	0.276	0.192	0.161
Observations	534	570	555	572	559	580	570
Panel B. Knows someone away							
Assigned to mentorship	0.183***	0.029	0.069*	0.045	0.041	0.027	0.019
	(0.070)	(0.018)	(0.035)	(0.028)	(0.035)	(0.030)	(0.029)
Control mean	0.006	0.910	0.484	0.786	0.376	0.232	0.195
Observations	765	821	802	811	806	823	804
p-value, knows someone = knows no one	0.537	0.873	0.074	0.918	0.979	0.965	0.991

Notes: Each column reports estimates from an OLS regression of the outcome on an indicator for assignment to the mentorship program, an indicator for knowing someone away from home at baseline, and their interaction. Robust standard errors in parentheses. Panel A reports the coefficient on assignment (knows no one away). Panel B reports the sum of the coefficients on assignment and the interaction (knows someone away). The last row reports the *p*-value of the interaction coefficient. Control means are for the control group of each group. *Knows someone away*: at baseline, the student knew someone (relatives, family friends, or friends) living in Pune, Mumbai, or Nagpur. The sample is all students with a completed endline interview; for students who could not be interviewed, outcomes are reported by a parent or another family member. All outcomes except the index are indicators equal to one if the student has taken the action or plans to within two months. Away from home means outside the student's home sub-district. *Applied to study*: applied for further study. *Applied to education financing*: applied for a scholarship or loan. *Enrolled in a program*: enrolled in a degree college, vocational institute, or coaching program. *Applied to study away*: applied to study away from home. *Migrated away*: moved away from home. *Migrated for education*: migrated away and enrolled. *Index of educational migration*: inverse-covariance-weighted average of the six components (Anderson, 2008), for students with all six observed, standardized to mean zero and standard deviation one in the control group. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table M3: Effects on educational migration, by decision-making power and gender

	Index of educational migration	Applied to study	Applied to education financing	Enrolled in a program	Applied to study away	Migrated away	Migrated for education
Panel A. Men, high decision power							
Assigned to mentorship	0.217*	0.088***	-0.014	0.038	0.114*	0.057	0.035
	(0.118)	(0.031)	(0.059)	(0.044)	(0.059)	(0.054)	(0.052)
Control mean	0.104	0.878	0.511	0.818	0.397	0.286	0.248
Observations	266	291	286	289	283	296	286
Panel B. Men, low decision power							
Assigned to mentorship	0.279**	-0.023	0.192***	-0.060	0.092	0.072	-0.013
	(0.134)	(0.038)	(0.064)	(0.048)	(0.064)	(0.058)	(0.056)
Control mean	0.021	0.916	0.398	0.867	0.400	0.248	0.250
Observations	228	241	240	239	237	243	238
Panel C. Women, high decision power							
Assigned to mentorship	0.126	0.018	-0.039	0.055	0.061	0.018	0.032
	(0.101)	(0.026)	(0.050)	(0.041)	(0.048)	(0.041)	(0.037)
Control mean	-0.013	0.916	0.588	0.754	0.315	0.199	0.143
Observations	381	406	394	402	400	406	400
Panel D. Women, low decision power							
Assigned to mentorship	0.071	0.035	0.037	0.095**	-0.049	-0.009	0.009
	(0.090)	(0.025)	(0.048)	(0.037)	(0.041)	(0.035)	(0.033)
Control mean	-0.065	0.909	0.526	0.759	0.279	0.170	0.138
Observations	424	453	437	453	445	458	450
p-value, men: low = high decision power	0.731	0.025	0.018	0.130	0.802	0.848	0.537
p-value, women: low = high decision power	0.682	0.637	0.270	0.466	0.084	0.625	0.648

Notes: Each column reports estimates from an OLS regression of the outcome on an indicator for assignment to the mentorship program, an indicator for low decision-making power, an indicator for women, and all their interactions. Robust standard errors in parentheses. Panel A reports the coefficient on assignment (men, high decision-making power). Panels B to D report the sum of the coefficient on assignment and its interactions with the indicators that apply to the group. The last two rows report the *p*-value of the difference between students with low and high decision-making power, for men and for women. Control means are for the control group of each group. *Decision-making power*: at baseline, students say who decides whether they continue in school after grade 12, whether they will work after their studies, what type of work they will do, which college they will attend, which city they will study in, and what they will study. The answers are "I decide by myself", "I decide together with my parents", and "My parents decide for me". Each answer is coded so that higher values mean more say and standardized using the control group, and the index is the average of the six standardized answers, standardized again using the control group. *Low decision-making power*: the index is below the median of all students at baseline. The sample is all students with a completed endline interview; for students who could not be interviewed, outcomes are reported by a parent or another family member. All outcomes except the index are indicators equal to one if the student has taken the action or plans to within two months. Away from home means outside the student's home sub-district. *Applied to study*: applied for further study. *Applied to education financing*: applied for a scholarship or loan. *Enrolled in a program*: enrolled in a degree college, vocational institute, or coaching program. *Applied to study away*: applied to study away from home. *Migrated away*: moved away from home. *Migrated for education*: migrated away and enrolled. *Index of educational migration*: inverse-covariance-weighted average of the six components (Anderson, 2008), for students with all six observed, standardized to mean zero and standard deviation one in the control group. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

M.2 Other pre-specified outcomes

Table M4 reports effects on willingness to migrate at midline, for all students and separately for men and women.

Table M5 reports effects on self-efficacy, gender attitudes and aspirations. The estimates are small and not significant.

Table M4: Effects on willingness to migrate (midline)

	Index of willingness to migrate	Probability of applying	Will look for colleges	Rating of destination	Choice experiment	College-rating exercise
Panel A. All students						
Assigned to mentorship × away	-0.017	0.025	-0.054	0.138	0.040	-0.080
	(0.079)	(0.021)	(0.039)	(0.181)	(0.085)	(0.203)
Control mean, away − home	-0.558	-0.120	-0.250	-0.704	-0.488	-0.421
Observations	1204	1204	1204	1204	1177	1204
Panel B. Men						
Assigned to mentorship × away	0.112	0.027	0.036	0.244	0.306**	0.231
	(0.116)	(0.030)	(0.064)	(0.263)	(0.129)	(0.322)
Control mean, away − home	-0.361	-0.063	-0.191	-0.333	-0.407	-0.270
Observations	430	430	430	430	418	430
Panel C. Women						
Assigned to mentorship × away	-0.096	0.022	-0.108**	0.065	-0.111	-0.258
	(0.102)	(0.027)	(0.048)	(0.238)	(0.112)	(0.258)
Control mean, away − home	-0.665	-0.152	-0.282	-0.904	-0.538	-0.506
Observations	774	774	774	774	759	774
p-value, women = men	0.179	0.913	0.072	0.615	0.015	0.237

Notes: Each column reports the coefficient on the interaction of an indicator for assignment to the mentorship program with an indicator for an option away from home. In the first four columns, each student contributes one observation for home and one for each destination away from home asked about, and the outcome is regressed by OLS on assignment, the away indicator, and their interaction. In the choice-experiment column, the estimate comes from a conditional logit of the chosen option on an indicator for a college away from home, indicators for the earnings level, an indicator for the opt-out option, and their interactions with assignment. In the college-rating column, the rating of each college profile is regressed by OLS on an indicator for a college away from home, its interaction with assignment, assignment, and the other attributes of the profile (tuition, living costs, earnings after graduation, share employed after graduation, admission cutoff, share selected for further study, and hostel). Panels B and C come from one regression per outcome in which every term is interacted with an indicator for being female. Panel B reports the interaction for men. Panel C reports the sum of that coefficient and its interaction with the female indicator (women). The last row reports the *p*-value of that last interaction. Standard errors clustered by student in parentheses. *Control mean, away − home*: the coefficient on the away indicator, the control group's answer about away destinations relative to home. The sample is all students interviewed at midline. Away from home means outside the student's home sub-district: the district capital, Pune, or Mumbai. *Probability of applying*: stated probability of applying to a college in the destination. *Will look for colleges*: plans to look for colleges in the destination; for the district capital, in another sub-district of the home district. *Rating of destination*: rating of the destination on a scale from 0 to 10. *Index of willingness to migrate*: inverse-covariance-weighted average of these three measures (Anderson, 2008), for students with all answers observed, standardized on control students' answers about home. *Choice experiment*: choices between hypothetical colleges at home and away that differ in expected earnings. *College-rating exercise*: ratings from 0 to 10 of hypothetical college profiles with randomly varied attributes. *Observations*: students. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table M5: Effects on self-efficacy, gender attitudes and aspirations

	Midline			Endline
	Self-efficacy	Gender attitudes	Aspirations	Gender attitudes
Assigned to mentorship	-0.050	-0.053	-0.033	-0.070
	(0.058)	(0.058)	(0.059)	(0.061)
Control mean	0.000	0.000	0.000	0.000
Observations	1204	1175	1204	1112

Notes: Each column reports the coefficient from an OLS regression of the outcome on an indicator for assignment to the mentorship program. Robust standard errors in parentheses. The midline columns use all students interviewed at midline. The endline column uses students interviewed directly at endline; proxy interviews did not ask these questions. Each outcome is an index standardized to mean zero and standard deviation one in the control group. *Self-efficacy*: sum of six statements of the shortened generalized self-efficacy scale, each rated from 1 (not at all true) to 4 (exactly true). *Gender attitudes*: average agreement with five statements, each signed so that higher values are more supportive of women's education: girls should be allowed to study in college even if it is far away; the student's parents think so; the student would discourage a sibling or cousin from studying in college far away; a university education is more important for a boy than for a girl; boys should get more opportunities for education than girls. At endline, only the first three statements were asked. *Aspirations*: average of standardized expected yearly earnings at age 25 (in logs), expected likelihood of working at age 25, and an indicator for wanting to live somewhere other than the current place at age 25. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.